1

Studies in theory and method in sociolinguistics

Volume ii

c John Pellowe

1990

Thesis submitted for the degree of Doctor of Philosophy in the University of Newcastle upon Tyne.

Appendix A

Establishing phonological criteria

1. Introductory considerations

First, I give definitions of the relevant components and units which constitute the form of criteria, together with a framework for interpreting the behaviour of such criteria.

Secondly, I discuss the reasoning which leads to the need for hierarchically defined criteria. The problem is deceptively simple: the framework of criteria must ensure that every sampled speaker is equally comparable with every other sampled speaker. The phrase 'every other sampled speaker' eschews the possibility of rejecting sampled speakers on the grounds of some a prioristic assumption of 'typicality' (which, anyway, begs theoretically important sociolinguistic questions)(A,C1). A community after all, whether it is to be considered from a linguistic or a sociological or a psychological point of view, must be as importantly characterised by its minorities as by its typicals - at least if it is to be recognisable as human.

Thirdly, I give a complete list of the particular criteria which represent the phonological subspace of the Variety Space which was introduced in general algebraic terms in Ch. 3. The particular criteria illustrate how the contents of that space are expressed at any particular moment. (This catalogue expands and annotates Pellowe et al. 1972a.)

Catalogues do not present themselves ready-made in the world of first-order percepts and first-order constructs, but are arrived at by a lengthy process of second-order construction, much of which may be, nevertheless, subconscious, or only semi-conscious (Kelly 1955, Bannister & Fransella 1971). This catalogue then, may be viewed as representing a series of psycho-archaeologically layered subjective approximations to, and thus deviations from, some optimal, computationally operable VSp. This catalogue is, I presume, only **one** possible picture of Tyneside's linguistic variety; and it has grown at different rates and with different delicacies at different points in its history. One's own growth of understanding of the nature of the structure of the linguistic variation in a community is, at least in part, a function of both one's perceived and one's attributed social roles and social relations **in** that community. To some extent then, this catalogue gestures towards a partial folk-sociolinguistic history of my own involvements in and on Tyneside.

2. Levels of criteria: (a) general definitions

It should be noted that all of the definitions which follow are non-essentialist definitions (sensu Janicki 1989,). That is, they do not purport to specify any essences, or ultimate realities, whatever they may be, in the manner of Aristotle. They simply constitute a set of more or less integrated shorthand labels for the purpose of attacking the problems specified in Chs. 2 & 3 above.

I State:

A symbol representing a phonetic realisation which is auditorily discriminable from all other different symbols will be called a state.

Notation: square brackets e.g. [i] [ʃ] etc

II Putative diasystemic variant (PDV):

A class of states which is sociolinguistically discriminable as a class from all other such classes, if any, within a particular Overall Unit (qv), such that the range of differences of the numbers of Lexical Sets (qv) and their composition in the population at large is adequately modelled will be called a putative diasystemic variant (= PDV).

[The second part of the definition ["such that . . ."] is an effective gloss for 'sociolinguistically discriminable'. Cf. Bierwisch (1977: 277) whose "pertinent difference[s] in connotation" are equivalent to such discriminations.] [N]

[N] Notice that exactly what might be meant by the evaluative term 'adequately' is not specified. It may be taken to imply that, in ideal circumstances, one would want to model the lexical sets in more than one way and then compare the effects of those different settings on a sample of speakers amongst whom the linguistic and social relationships were already very well understood. But it seems clear that circumstances will never be ideal and that the universe of comparisons is, anyway, transfinite or ℵ, in Cantor's scheme of things (Langendoen & Postal 1984) (A,C2).

Notation: double slants e.g. *//i//* *//iə//* *//eɪ//* etc.

III Overall unit (OU):

An arbitrarily chosen abstract (phonological) symbol which encapsulates the complete lexical set in which it occurs will be called an overall unit (OU).

[The set of all OUs thus totally expresses one possible full series of lexical sets. But it is not necessarily the case that that full series is an effective representation of the lexical sets of any particular, actual, single speaker.]

[Though the the symbol associated with an OU is an abstract phonological one, it so happens that the OUs specified here are more closely related, but not completely so, to a system of non-localised British English than to a localised system. But this is of no theoretical importance: in principle OUs could be based on any system whatever, provided only that all the analysts involved knew unambiguously to which lexical set to assign some newly sampled word in respect of its **n**th segment. (In respect of each of its **n** segments, every word is in one and only one OU.) Here, the particular system to which an OU is related is indicated by superscript, thus: $\{\}^{NL}$ or $\{\}^{L}$. Regardless, therefore, of what phonetic realisation represents a segment in a word, we must be able to determine the unique lexical set membership of any word carrying that segment, <u>in terms of the complete set of OUs</u>. The OU thus ensures comparability between varieties <u>whatever their degree of sameness or difference</u>.]

Notation: curly brackets e.g. $\{e\iota\}^{NL}$ $\{a\}^{L}$ etc.

IV <u>Lexical set:</u>

A list of institutional words which is, in principle, totally enumerable will be called a <u>lexical set</u>.

Notation: curly brackets e.g. {see, we, knee, be, . . .}

These definitions (I - IV) may be exemplified as follows.

Lexical set: {wall, fork, port, paw, war, all, walk, auction, horse, more, sore, four, door, course, . . . }

Overall unit: {ɔ}NL

Putative diasystemic variants
(PDVs of OU {ɔ}):

1. //a//	{walk, wall, . . . }	[5 states]
2. //ɔ//	{fork, war, . . . }	[7 states]
3. //ɒ//	{auction, horse, . . . }	[5 states]
4. //ɛ//	{more, sore, . . . }	[3 states]
5. //aɷə//	{four, more, . . . }	[2 states]
6. //ɷə//	{door, course, . . . }	[4 states]

(The relevant states for each PDV of each OU are presented below.)

V <u>Quantitative (criterion)</u>:

A variable about whose systematic relationships we know too little to be able to place it in a hierarchical qualitative structure (of the type in which PDVs appear) will be called a <u>quantitative (criterion)</u>.

[Quantitatives operate over an interval rather than a nominal scale; that is, they each may assume any one of a number of continuous arithmetic values, as opposed to any one of an unordered set of discrete states. In this work, prosodic and paralinguistic features are treated as quantitatives - see below, Appendices B & C & D, for more details of the problems and methods of handling them.]

VI **Profile**:

A representation of the utterances of a single speaker on a single occasion which is exhaustive in terms of PDVs and their states, and of quantitatives, will be called a profile. [N]

[N] I use the term 'analytic variety' or ${}_{a}V$ under the same definition when I wish to focus upon its being a point in an n-space rather than a list of values of the criteria. See above, Ch. 3, for discussion of the distinction between ${}_{a}V$ and ${}_{h}V$.

VII **Variety space (VSp)**:

A multi-dimensional space comprising a mixture of nominal, ordinal and interval variables (linguistic criteria) which are assumed, initially, to be orthogonal or independent will be called a Variety Space (VSp).

VIII Variety cluster (VC):

A number of analytic varieties ($_aVs$) which show some degree of association in response to a comparison function traversing the variety space will be called a Variety Cluster (VC).

IX Diagnostic:

Some value of any given variable which may be interpreted as especially characteristic of a particular VC will be called a diagnostic.

[Note that the manner in which the variable is seen as being 'especially characteristic' must be **chosen.**]

[A diagnostic is thus not, strictly, a property of the VSp itself, but is a property of a subset of the contents of the VSp as they are dispersed on the axes of the VSp and associated by the comparison function.]

X Diagnostic profile:

A listing of all diagnostics for a particular VC will be called its diagnostic profile.

XI <u>Key diagnostic</u>:

A member of a subset of the diagnostic profile of a particular VC, for which some well-defined measure of predictiveness or discrimination is high will be called a <u>key diagnostic</u>.

[Note, again, that the measure must be **chosen**.]

XII <u>Localised variety</u>:

Given a random sample of some population of speakers, and given some suitable method of measuring similarity amongst those speakers and forming groups of them on the basis of that similarity (see below, Appendix D), then at least one part of the space (VSp) will be, generally, more densely filled with VCs than another; the component $_{a}$Vs of such VCs (and hence the VCs themselves) will be called 'localised'.

[Such VCs are not only 'localised' in the VSp, but their members have a tendency to be geographically locateable (with varying degrees of precision) by at least some hearers.

Conversely, VCs having diagnostic profiles in common may occur in geographically quite disparate samples (if those samples are big enough to overcome the problem of rarity, see above Chapter 3), and may thus be labelled 'non-localised'. From another point of view, we may conjecture that the probability of association between non-localised varieties on the one hand and a whole range of different localised varieties on the other hand will be roughly the same in each case.]

2. Levels of criteria: (b) rationale

General computational reasons for the form of the criteria are given below in Appendix D. Here, by contrast, I am concerned only with setting out general sociolinguistic arguments in favour of the definitions of units, levels and distinctions which are given in detail above. [N]

[N] Within a strict non-essentialist framework the set of definitions which precedes renders the rationale here redundant, though not perhaps the exemplification. Indeed, insofar as a reading of the rationale assumed that its distinctions were believed by the writer to be components of a necessary truth, just so far one would have to assert that that rationale was misleading. Nevertheless, it is included here since its matter reflects early discussions between Barbara Strang, Vince McNeany and myself, and it is therefore a historically important part of the T.L.S.

The reasoning which follows arises from some problems discussed by Pellowe (1967; 1970 a,b,c; Pellowe et al. 1972) and draws on the methodology of the diasystem (particularly Weinreich 1954, Moulton 1960, Pulgram 1964).

In Chs. 2 & 3 above I argued for the view that an appropriate model of the hearer's processing of speech variation must incorporate mechanisms for handling varieties which had not previously been encountered by that hearer. I suggested that the consequence of this view for the model was that it would not be allowed to discard _any_ speaker who was sampled in order to be located in its Variety Space. This requirement raises the issue of the comparison of, and the comparability between, profiles of linguistic variants.

Weinreich's goal was to secure a useable _rapprochement_ between dialectological concerns and structuralist principles, but he acknowledged that this would be only possible if structural linguistics abandoned the illusion of a perfect system in which _tout se tient_, since in such a framework there could never be any chance of admitting that two systems could be _partially_ different (or _partially_ similar) - systems could be either identical or _incommensurably_ different (simply distinct). Given a willingness to admit the possibility of partial similarities and differences between different systems, what is needed are

"procedures for constructing systems of a higher level out of the discrete and homogeneous systems that are derived from description and that represent each a unique formal organisation of the substance of expression and content" (1954: 389-90)

Such constructions Weinreich calls diasystems. He points out that a diasystem can be constructed "out of any two [or more] systems which have partial similarities" (1954: 390). It is a significant element of support for the theory presented in Chs. 2 & 3 that Weinreich also says (1954: 390) that such constructions are not necessarily only those of the scientist but may well also be the inferential basis for information derived by any single '**ordinary**' speaker-hearer.

In constructing a diasystem, it is important to distinguish between differences of inventory and differences of distribution. Weinreich (1954: 394) gives four examples of diasystemic inventories. (Varieties have subscript numbers, single slashes (/) enclose sets of phonemes, single inequalities (<>) symbolise systemic opposition, double inequalities (<<>>) symbolize oppositions/contrasts in the diasystem, double slashes (// //)symbolise the boundaries of the diasystem.)

(a) A diasystem of two varieties with identical five vowel systems:

$_{1,2}$ // i <<>> e <<>> a <<>> o <<>> u //

(b) A diasystem of two varieties, in the second of which the second vowel is always more open than it is in the first:

$$_{1,2}//\ i \langle\langle\rangle\rangle \frac{_1\, e}{_2\, \varepsilon} \langle\langle\rangle\rangle\ a \langle\langle\rangle\rangle\ o \langle\langle\rangle\rangle\ u\ //$$

(3) A diasystem of two varieties, in which the first has three front vowels whilst the second has four:

$$_{1,2}//\ \frac{_1\, /\, i \langle\rangle e \langle\rangle æ\, /}{_2\, /\, i \langle\rangle e \langle\rangle \varepsilon \langle\rangle æ\, /} \langle\langle\rangle\rangle\ a \langle\langle\rangle\rangle\ o \langle\langle\rangle\rangle\ u\ //$$

(4) A diasystem of three varieties of Yiddish (1 Central, 2 Southwestern, 3 Northwestern):

$$_{1,2,3}//\ \begin{array}{c} _1\, /\, iː \langle\rangle i\, / \\ \hline _2\, /\, i \langle\rangle ɩ\, / \\ \hline _3\, i \end{array} \langle\langle\rangle\rangle\ e \langle\langle\rangle\rangle \frac{_1\, /\, aː \langle\rangle a\, /}{_{2,3}\, a} \langle\langle\rangle\rangle\ o \langle\langle\rangle\rangle\ u\ //$$

We see then that in order to construct a diasystem, _sensu_ Weinreich, we must place discrete, well-defined varieties in a kind of continuum determined by their partial similarities. Weinreich insists that a clear picture of differences in inventory between the varieties is a _prerequisite_ for establishing a diasystem (1954: 394). We shall see below that one of the ways in which Weinreich's notion has to be adapted to our purpose is as a result of our not being able to state the differences of inventory between the varieties at the outset. It is also important to emphasise (and see also Moulton's (1960) remarks below) that "differences in distribution cannot be directly inferred from a comparison of the differences in inventory" (1954: 394).

Regrettably, apart from the few remarks cited above, Weinreich does not explore either the relationship between items in the inventory and the (lexical) distribution of those items, or the problems which arise from that relationship. Cochrane (1959) in a diasystemic treatment of vowels in Australian English does face up to that relationship. He believes that the diasystem must be set up without reference to the lexical distributions of its elements in the contributing varieties; and that these distributions must be dealt with only subsequently. However, Moulton (1960: 176) rightly, in my view, emphasises that this is neither as simple nor as revealing as it seems.

He points out that if a diasystem is constructed for two varieties of a language without consideration of distributional matters, then that is tantamount to treating the two varieties as totally unrelated,

<u>even though they may be identical in the diasystemic representation.</u> Alternatively, if a diasystem is constructed taking due account of the lexical distribution of the elements of its contributory varieties, then though the varieties are thereby being taken as related, the diasystemic account may show them as far more <u>dissimilar</u> than is reasonable or desirable (whether from the point of view either of the dialectologist or of the native variety speakers themselves). Moulton (1960: 176) illustrates with two varieties of Swiss German (LU(zern) and AP(penzell)):

if lexical correspondences are disregarded -

LU,AP//i <<>> e <<>> ε <<>> æ <<>> a <<>> ɔ <<>> o <<>> u <<>> ü <<>> ō <<>> ɔ̄//

which makes the varieties too similar;

if lexical correspondences are included -

$$\text{LU,AP//}\ \frac{\text{LU/}i_0 \langle\rangle e_1 \langle\rangle \varepsilon_2 \langle\rangle æ_{3,4}/}{\text{AP/}i_{0,1} \langle\rangle e_{1,2} \langle\rangle \varepsilon_3 \langle\rangle æ_4/}\ \langle\langle\rangle\rangle\ a_4\ \langle\langle\rangle\rangle\ \frac{\text{LU/}ɔ_2 \langle\rangle o_1 \langle\rangle u_0 \langle\rangle ü_0 \langle\rangle \bar{o}_1 \langle\rangle \bar{ɔ}_2/}{\text{AP/}ɔ_2 \langle\rangle o_{1,2} \langle\rangle u_{0,1} \langle\rangle ü_{0,1} \langle\rangle \bar{o}_{1,2} \langle\rangle \bar{ɔ}_2/}$$

which makes the varieties too dissimilar.

(Moulton's subscripts here refer to historically earlier lexical sets. Thus the set of lexical items associated with /e/ in LU is distributed across both /i/ and /e/ in AP; the two sets of lexical items realised by /ε/ and /æ/ in AP are both realised by /æ/ in LU; LU/$ɔ_2$/ is not the same in lexical distribution as $AP/ɔ/_2$, since part of the lexical set realised by $LU/ɔ/_2$ is realised by $AP/o/_{1,2}$; and so on, and so on.)

The problem which he highlights here was not solved by Moulton. We may as well call it Moulton's Paradox:

> when, for **n** partially similar varieties, a diasystem is constructed excluding the facts of lexical distribution, that diasystem represents the **n** varieties as too similar; conversely, when, for **n** partially similar varieties, a diasystem is constructed including the facts of lexical distribution, the **n** varieties are too dissimilar.

My adaptation of the diasystemic method for the purpose of representing a multiplicity of urban speech varieties whose systems cannot be (wholly) known in advance shows that Moulton's Paradox can, in fact, be dissolved, even if at the price of a certain amount of induced (and specifiable) redundancy. My dissolution of the problem is given in abstract definitional form above. Here I shall merely give a discursive account and conclude with some illustration.

We need to be able to compare the realisations of all speakers in an entirely systematic way. Consider the following realisations (diacritics are omitted):

Form	neat	treated	bread
Speaker 1	[niːt]	[tɹiːtɪd]	[bɹɛd]
Speaker 2	[niːt]	[tɹɛt]	[bɹiːd]
Speaker 3	[nɪət]	[tɹiːɾəd]	[bɹɛəd]

First, in order to ensure undistorted comparability, we must find a way of coding the different realisations as variants of something. A given segment in a given lexical item must always be coded as a variant of the same entity, regardless of the rarity or oddity of its realisation. This 'something', this 'same entity' is established as a series of fixed lexical sets, which, in principle exhausts the lexicon of English. Each lexical set is that associated with its Overall Unit. Then, regardless of what the phonetic realisation of some particular phoneme is, and regardless of what system that phoneme is a member of, we shall always know within which lexical set comparisons between that phoneme and phonemes of other, more or less different systems must be made. Thus there is a lexical set

{week, relief, see, field, beat, treat, neat, sea, we, . . . }

which is associated with $_{NL}$/i:/. We set up an OU {i:}^{NL} which expresses this set. Another OU {ε}^{NL} will express the lexical set

{head, bread, centre, never, engine, many, well, fell, men, . . .}.

Then **all** realisations of the vowel of neat, treated, will be coded under OU {i:}, and **all** realisations of the vowel of bread will be coded under OU {ε}.

But it is clear from the forms transcribed above that different speakers distribute subsets of the OU-lexical set in different ways – whether across the 'same' speech sounds in different subsets, or across more or fewer different speech sounds. And it is clear that individual speakers and groups of speakers do this consistently.

We shall associate with each of these consistent different patternings of the lexical set a putative diasystemic variant (PDV). This name simultaneously asserts three things:

(a) that, contrary to Weinreich's prerequisite, we do not know the complete phonemic inventory of every distinct speech variety on Tyneside before we analyse data;

(b) that the phonological status of the symbols associated with these consistent different lexical subsets is phoneme-like rather than phone-like; and

(c) that until larger samples of varieties from the same population show that the set of PDVs is both necessary and sufficient to express those varieties, these diasystemic variants can only be putative.

Returning to the above example, we note that the OUs {iː} and {ɛ} will have amongst their PDVs the following:

OU {iː} PDVs *//iː//, //ɛ//, //ɪə//*

OU {ɛ} PDVs *//iː//, //ɛ//, //ɛə//*.

Then speakers 1, 2, and 3 above may be represented as follows:

	OU {iː}			OU {ɛ}		
	PDVs *//iː//*	*//ɛ//*	*//ɪə//*	PDVs *//iː//*	*//ɛ//*	*//ɛə//*
Sp.1	+	–	–	–	+	–
Sp.2	+	+	–	+	–	–
Sp.3	+	–	+	–	–	+

This manner of representation - by means of OUs and PDVs - clearly shows the similarities and differences between the speakers in terms of their unique partitionings of the lexical sets of the OUs. (Notice that the method leaves open the question of whether, for a speaker who had [bɹɛd] [tɹɛt] in bread, treat, the two [ɛ] realisations are realisations of one phoneme or of two, though this particular implementation of the OU/PDV system inclines one to imagine the latter.)

Within any given PDV, there will be several possible phonetic realisations (States). There are thus two levels upon which speakers (varieties) may be specified as more or less similar to one another: PDVs and Sts. Two speakers having the same PDV and the same St in some segment of some word are thus 100% similar; two speakers having the same PDV and different Sts are 50% similar; two speakers having different PDVs are 0% similar (clearly, in the latter case, the issue of Sts does not arise). (Notice that the OU level contributes nothing to the assessment of the overall similarity: it is simply a device to ensure undistorted comparability across varieties.)

It is important to stress that PDVs are not based on phonetic difference alone, but in addition, on the emergent psychological reality which arises from the coherence attributed to a speaker's lexical subsets. (Note "sociolinguistically discriminable" in the definition of PDV above.) If it <u>were</u> the case that PDVs were only based on phonetic difference, it could be argued that a PDV level of organisation was redundant, since one could capture the same differences by proposing to compare speakers on the basis of a free list of phonetic states for all possible lexical subsets of an OU. Because PDVs are <u>not</u> based on phonetic difference alone, it can be the case that two different PDVs <u>of the same OU</u> may have one or more Sts in common.

Where, it may be asked, do these PDVs come from ? They are observational constructions from many spoken sources over a period of years. (The sources were: overheard talk in public places, surreptitious recordings, notes taken of speakers on "Voice of the North" (BBC Newcastle, daily), miscellaneous tapes in the TLS archive.)

A variant was admitted as a PDV if:

(i) it had been heard used in items of the same fuzzily coherent lexical subset by at least five different speakers under stylistically comparable conditions; and if

(ii) it had been recognised by at least one localised hearer as a local realisation.

Condition (i) does not demand that the variant's potential for occurrence in the lexical set be known in complete detail - such an impossibility is one of the reasons we cannot satisfy Weinreich's prerequisite of establishing complete inventories before doing anything else.

Condition (ii) acknowledges that something like an open-ended diasystem must be the probabilistic basis for a lot of hearer inferencing about the meaning of speaker variations.

It is also important to point out that the establishment of a PDV is in no way based on frequentist notions. It is reasonable to set up a PDV for a very small lexical subset of the OU, in the limiting case even only of one lexical item. For instance the PDV //ιɛ// in the OU {əω} only has one lexical item [ɕιɛm] [jɛm] <u>home</u>. The PDV //ɛ//of the OU {eι} only has two well-attested lexical items - [mɛk] [tɛk] for make, take - but may be permeable to analogy. (See below Section 3.(a) for details.)

An interesting application of the above definitions and methods to what otherwise might look like a 'traditional' dialect study of Grassington (North Yorkshire) speech may be found in Glauser (1984).[N]

[N] Closer inspection in respect of matters other than the above reveals just how untraditional a study this is, since it presents, though all too briefly, analyses of tone, tempo, rhythmicality and pause, and their cooccurrence in a sample of running text (following Crystal & Quirk 1964; Crystal 1969).

Here is another example of a defined criterion, illustrating how the conventions for coding particular speakers are applied, and the extent to which the definition of criteria depends upon aspects of comparability.

OU	No.	PDV	Examples *	Sts
(i)	1.1	//i//	feed, week, relief, treat(ed), beat(en), see, . . .	1 to 6
	1.2	//ɪ//	week, relief, . . .	1 to 5
	1.3	//ɛ//	treat(ed), beat(en)	1 to 4
	1.4	//eɪ//	see, feed, . . .	1 to 4
	1.5	//ɪə//	relief, feed, . . .	1 to 3
	1.6	//ɪi//	knee, see, . . .	1 to 3

(Thus there are six Sts of //i// which we wish to consider, there are five of //ɩ//, and so on. The number of each state is only its label within a particular PDV; there is no implicit ranking and it is not an identification; i.e. St 4 for PDV 1.2 is not the same as St 4 for PDV 1.4. See below, Sections 3, 4, & 5, for details of Sts.)

* It is of critical importance to understand that relevant exemplars of PDVs can differ from one variety to the next. That is, two different varieties, **a**, **b**, may have, in the single OU {i}, the **same** PDVs - e.g. 1.1, 1.4, & 1.6 - but the lexical subsets associated with each PDV in the two varieties **may be different**.

For most PDVs of most OUs one has a fairly good idea of the size of the lexical subsets associated with the PDVs - for certain kinds of Speaker. One may therefore speak, with differing degrees of confidence, about partitions of the lexical set associated with some given OU. A partition of a particular lexical set is thus the set of probabilities of particular duples {lexical $item_i$: PDV_j} in a given variety. It can be no more than a set of probabilities because all possible members of the partition (or the complement of the partition) cannot possibly occur in any interview. The notion of a majority partition follows naturally. A majority partition is one in which the majority of lexical items in the lexical set (OU) is realised by one PDV.

In some cases this will be true of all varieties, whether localised or non-localised; in some cases there will be two competing majority partitions - one for localised varieties, one for non-localised; in some cases there will be two competing majority partitions for localised varieties; and so on.

Notice that the definitions of OU, PDV, and the notion of partition do not, <u>in any manner</u>, constrain or prejudge the ways in which any particular variety may express, combine, borrow or overlap its particular partitions of particular lexical sets (A,C3). The fact that the same lexical item may exemplify different PDVs of the **same** OU (see Table above) is intended, by reflecting actual distributions in the speech community, to give substance to this claim.

The table above gives six qualitative multistate (vowel) criteria (1.1 to 1.6). Potentially, a speaker may be scored positively on all six PDVs, or on only one.

Given that all the occurrences of this OU, {i}, are realised by one PDV,

//i//, in the speech of a particular speaker, s, then she is only permitted to score on one state, namely the numerically dominant one. This is an intentional, but not a necessary constraint. What both it and the definitions exemplified in the table assume is that the PDVs should be weighted. And they should be weighted

(a) not less than the states, which would be the result of coding more than one state per PDV, and

(b) not as a function of the number of states each PDV has, since this is open to manipulation or chance definition,

but they should be weighted indirectly, as a function simply of the number of PDVs which have been observationally established for each overall unit. [N]

[N] They are indirectly weighted because although each PDV has, for the OU illustrated, an a priori 1/6 probability of turning up compared to an OU with only one PDV, nevertheless the group (VC) forming power and the diagnostic potential of each is 6x greater. Since it is the VCs and the diagnostics which are of interest, this weighting method is important.

This weighting again brings us back to the hearer, since if more

PDVs have been observed for one OU than for another, then they probably **do** differentiate more for the hearer in the former case than in the latter, and they probably therefore **should** contribute more to the speaker's profile. In other words this method is one which prevents phonetic variation from overwhelming phonemic or near-phonemic variation.

Thus for speaker **x** we have the following profile (positive responses only; parenthesized numerals refer to number of occurrences; → = 'realised by')

	OU		PDVs		Sts
x =	{i:} (18)	→	//i// (12)	→	[i] (4)*
				→	[i̠̞] (8)
		→	//ɩi// (6)	→	[əi̠̞] (5)
				→	[ɨi̠̞] (1)*

Sts marked * are supressed because only the dominant St is scored for any given PDV, so more briefly,

x = {i} → //i// ↔ //ɩi// → [i̠̞] ↔ [əi̠̞]

where '↔' means 'is in co-distribution with'. (For reasons which should be by now clear, I wish to avoid the words 'co-variation' and 'alternation'.) The order of the dominant Sts is directly associated with the order of the PDVs.

And, in the same notation,

y = {i} → //i// ↔ //ɛ// ↔ //ɩə// → [i̞] ↔ [e̱] ↔ [i̞ɛ̱.˞]

z = {i} → //i// ↔ //ɩə// ↔ //ɩi// → [i] ↔ [i̞ɛ̱.] ↔ [i̞i̞]

All vowels and consonants are defined in this way, as qualitative multistate criteria. The profile for any particular speaker, **x**, is a long list of strings similar to that for {i}, above. But the actual comparison of speakers on the basis of their profiles is not straightforward.

First, I have already mentioned (above, Ch. 2) that some hearers, under certain conditions, diagnose some varieties, at least partially, by means of the **absence**, rather than by means of the presence, of variants. This being the case, we shall have just as much need of the <u>negative</u> diagnostics of a VC as of its positive diagnostics. And if these negative diagnostics are to reflect regularities in the data which we wish to examine (rather than merely reflecting, for instance, glitches in the data collection, or the infinite dark universe of the unconsidered and the unimagined), then our comparison functions must incorporate negative matches as part of what is similar between varieties. [N]

[N] Sneath (1957b) rejects negative matches on the grounds that the universe of negative criteria is infinite. But then the universe of positive criteria is also infinite; according to Johnson (1970: 213) "every object has an infinitude of attributes". Carvell & Svartvik (1969) claim that negative matches are less information-bearing, in predictive uses of classification, than positive matches. But clearly such a claim rests primarily on the definition of 'information', rather than on the nature of negative matches, and even if this criticism is set aside, what is being complained of in their claim is less likely to be a property of the data, than a consequence of either the form of definition of the criteria, or the form of sampling of the criteria, or the subordination of the selected criteria to criteria which are relevant to some hierarchically higher level of data structure. MacNaughton-Smith (1965: 10), in a wide ranging review of taximetric techniques likely to be useful in the human sciences, concludes that "it is doubtful whether [omitting negative matches] should ever really be done".

Secondly, if, by some extraordinary chance, representative items of an OU such as {i} did not arise in all of the taped speech of some informant, then there could by definition be no instances of PDVs 1.1 to 1.6. To mark them as negative would be misleading since it is not so much that they are absent, but unable to be present. Such cases are coded NC. The same principle is applied to its Sts when a particular

PDV is negative.

To illustrate, consider again the partial positive profiles of speakers **x**, **y**, and z.

I tabulate their profiles separately on PDVs and on Sts.

OU	PDVs	Speakers			Possible Sts	St realised [N] by speaker		
{i}		x	y	z		x	y	z
	//i//	+	+	+	(6)	6	6	2
	//ɩ//	-	-	-	(5)	NC	NC	NC
	//ɛ//	-	+	-	(4)	NC	1	NC
	//eɩ//	-	-	-	(4)	NC	NC	NC
	//ɩə//	-	+	+	(3)	NC	2	2
	//ɩi//	+	-	+	(3)	2	NC	3

[N] The numbers in the body of this table are labels: there is no ordering principle involved amongst states of one PDV, since they are qualitatives. Thus Minkoff's (1965) strictures - that the ordering of states of (quantitative) criteria radically affects the value of the comparison function (e.g. similarity coefficient) - do not apply.

When we compare these three speakers by pairs, recalling that

(a) negative matches count as similarities, and

(b) if either member is scored NC we do not count a similarity, we get:

xy	:	PDVs	comparisons:	6
			matches:	3
		Sts	comparisons:	1
			matches:	1

then the similarity between **x** and **y** is

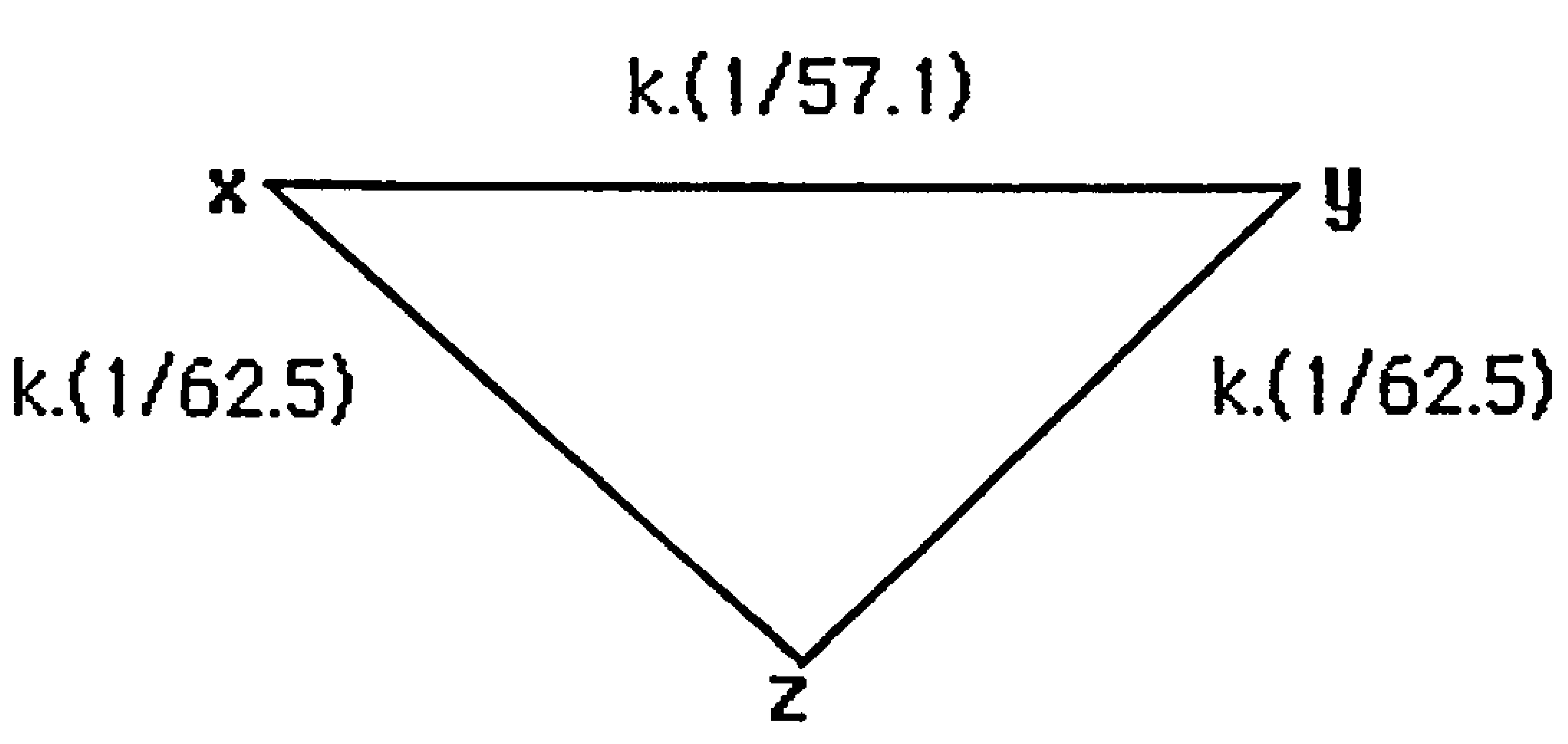
$$sim_{xy} = matches/comparisons = (3+1)/(6+1) = .571 = 57.1\%.$$

Similarly, but more briefly,

$$sim_{xz} = 5/8 = 62.5\%, \text{ and}$$

$$sim_{zy} = 5/8 = 62.5\%.$$

Diagrammatically, distance being some function (k) of the reciprocal of similarity, we can represent this as



This is a two-dimensional representation of a truncated sub-space of the Variety Space. (It is a truncated subspace because criteria 1.1 to

1.6 represent only a fraction of those concerned with vowels in stressed syllables.)

Notice firstly, that the product of the comparison function - i.e. similarity or its inverse in the present case - is only an approximating shorthand for, in the present example, six Cartesian coordinates; it is approximating because though **x** and **y** have the same distance function from **z**, they do not have the same coordinates as each other. Secondly, from the diagram above it is not possible to state whether the similarities between the speakers are, per se, within group or across group relationships. (Such statements depend jointly on the range of the comparison function in the whole sample, and on the method of forming VCs. See further below, Section 7 and Appx. D.)

3. Vowels in stressed syllables:
(a) criteria for vowels in stressed syllables

OU {i:} NL {week, relief, treat, beat, see, feed, field, sea, we, . . }

The lexical set subsumed here is that usually represented by /i:/ for 'RP'. The superscript NL (non-localised) indicates that this is the case.
The numbering of the PDVs represents a coding system now of no interest or significance. It is retained only to facilitate comparison with Jones (1978) and Pellowe et al.(1972a); the numbering follows the latter very closely. Any particular state can be expressed by a five digit code, four digits for the PDV and one for the relevant state of that PDV. Thus a code of 00021 indicates a segment which is phonetically Cardinal Vowel 1, represented as [i], occurring in PDV //i// of an item from the lexical set associated with OU {i} NL.

The numbering of the criteria by Jones (1978 [=Jones-Sargent 1983]) in her Appendix A, pp. 294 ff., is the **same** as that given here. However, her program for accumulating the sums of tokens (VAL 1) [Jones 1978: 78 ff.] simply gives, for each successive triple, {OU: PDV: State}, - or duple, {OU: State} - a serial number from 1 to 690.

0002 PDV //i:// (the majority partition)
states (6): i i̞ i̠ i̠̞ ɪ i̠̞̞ (= ẹ)
Cf. all e.g.s of OU (above)

0004 PDV //ɩ//
states (5): i i̞ i̠ i̠̞ ɪ
e.g. week, relief

0006 PDV //ɛ //
states (4): e̞ ɛ̞ e̞˕ ɛ.
e.g. treat(ed), beat(en)

As far as I know these examples exhaust the partition – i.e. speakers of varieties in which these words are pronounced in this way do not pronounce <u>any</u> other items from the lexical set in this way. The PDV occurs not only in preterite and past participle forms but also in the base infinitive following periphrastic (but rarely emphatic) <u>do</u>. Whenever the PDV occurs, the <u>-ed</u>/<u>-en</u> morphemes are **absent**, i.e. realised by a Ø morph.

0008 PDV //ei //
states (4): e̞.i ə˃i e̞˕˃ ɛ̞i̞˕
e.g. see, feed

0010 PDV //ɩə//
states (3): i̞˕˕ɛ̞.˃ i̞˕˕ɛ̞. i̞˕ə
e.g. feed, field

0012 PDV //ɩi //
states (3): ɩi [back] ɩi [low] əi̞˕
e.g. we, see

OU {ɩ} NL {live (vb.), fit, it, big, till, . . . }

0014 PDV //ɩ// (the majority partition)
states (5): i̞ i̞˕˕ i̠ i̠˕˕ i
Cf. all e.g.s of OU

0016 PDV //ɐ//
states (6): ə ə̞̞ ə̞̞̞ ə̝ ɜ̠‹ ɜ̠.‹
e.g. shilling, till

0018 PDV //ɩə//
states (3): i̞̞ɛ̠.› i̞̞ ɛ̠. i̠̞ ə
e.g. did, live (vb.)

0020 PDV //ɜː//
states (3): ɜ̠ ɜ̠. ɛ̠.
e.g. live (vb.), shilling

0022 PDV //ɛə//
states (2): e̠ ɨ ɛ̠ ɨ
e.g. live (vb.), Miss, big

OU {ɛ} NL {head, bread, centre, never, engine, many, well, fell, men, . . . }

0024 PDV //ɛ // (the majority partition)
states (6): ɛ̞ ɛ̠ ɛə ɛ. e̞ ɛ
Cf. e.g.s for OU

0026 PDV //iː//
states (5): ᵊi̠̞ i̠ i̠̞̞ ɨ i̠̞ᵊ
e.g. head, bread

0028 PDV //ɩ //
states (5): ɨ i̠ i̞̞ i̠ i̠̞̞
e.g. centre, never

0030 PDV //ə//
states (5): ə ə› ə̤› ə̤ ə̣.
well, <u>man</u>y

0032 PDV //εə//
states (2): εə ẹ(ə)
men, <u>em</u>bassy

OU {æ} L {path, grass, have, <u>al</u>sation, pal, <u>mas</u>ter, <u>lath</u>er, . . .}

With this and the next two OUs I am giving effective voice to the claim in the definition of OU that they are arbitrary abstract symbols whose selection does not <u>represent</u> any set of real phonological contrasts in some actual variety of English.
Basing these OUs on in-principle definable lexical sets of Northern British English shows that the normalising effect – to allow PDVs to interact dynamically – is achieved without distorting the classificatory procedures.

0034 PDV //æ//(the majority partition)
states (5): a̠. a̠. ạ ạ. ā
Cf. all e.g.s of the OU

0036 PDV //ε//
states (4): ə̣ ə̤ ε̤ ε̣
e.g. have, after

0038 PDV //ɑ//
states (5): ɑ̣ ɑ̣› ɑ̤ ɑ ɑ̣
path, grass

0040 PDV //ɔ//
states (7): ɔ̣˔ ɒ ɒ̣ ɔ̣ ɔ̣˔˂ᵊ ɔ̣˔˂ ɔ̣
alsation, pal

OU {ɑ} L {father, barn, farm, card, half, bar, rather, are, . . .}

0042 PDV //ɑ// (the majority partition)
states (5): ɑ̣˔ ɑ̣˔˃ ɑ̣˔˔ ɑ ɑ̣
Cf. e.g.s for the OU

0044 PDV //ɔ//
states (7): ɔ̣˔ ɒ ɒ̣ ɔ̣. ɔ̣˔˂ᵊ ɔ̣˔˂ ɔ̣
e.g. farm, card

0046 PDV //æ//
states (4) a̠. a̠̣ ạ ạ.
e.g. father, half, rather

OU {ɒ} L {off, often, one, involved, along, watch, holiday, swan, because, wash, long, . . .}

0050 PDV //ɒ// (the majority partition)
states (5) ɒ̣ ɒ̣˔˔˃ ɒ̣˂ ɒ̣˔ ɒ
Cf. e.g.s for the OU

0052 PDV //ɔ//
states (6): ɔ ɔ̣.˔ ɔ̣˃ ɔ̣ ɔ̣˔˔ ɔ̣
e.g. off, often, involved

0054 PDV //a//
states (5): a̠ a̠̣. ạ ɛ̠ ɛ̠˕
e.g. along, swan, watch, holiday

0056 PDV //ə//
states (5): ə ə̤˃ ə̤ ə̣ ə̤
e.g. be<u>cause</u>, curi<u>o</u>sity
[partition not much more extensive than these examples]

0058 PDV //ɛ//
states (2): ɛ̱ e̞
e.g. wash, long, watch

5800 PDV //ʌ//
states (4): ɔ̤ ɔ̤̣ ɑ̣ ã
e.g. one, none

5600 PDV //ɷ//
states (5): ṳ˂ ọ˂ o̤˂ o̤˂ ṳ˂
e.g. be<u>cause</u>, once

OU {ɔ} NL {all, war, talk, or, four, more, door, course, auction, . . . }

0060 PDV //a// (large localised partition, but excludes orthographic <u>-or</u>, <u>-our</u>)
states (5): a̱ a̩ a. ɛ̱ ɛ̞
e.g. all, talk, war

0062 PDV //ɔ// (the majority partition)
states (7): ɔ ɔ̤ ɔ̣˃ ɔ̞ ɔ̤ ɔ̣ ɔ̞˂
Cf. e.g.s for OU

0064 PDV //ɒ//
states (5): ɒ̣ ɒ̤˒ ɒ̣˓ ɒ̣.˓ ɒ
e.g. auction, horse

0066 PDV //ɛ//
states (3): ɛ̣ ɛ̣ə ɛ̣.
e.g. more, sore

0068 PDV //aωə//
states (2): ăωə ɑ̣.ωə
e.g. four, more

0070 PDV //ωə//
states (4): ω̣.(ə) ωə̠ (ə̠)ω̣ə (ə̤)ω̣ə
e.g. door, course

OU {ʌ} NL {hurry, onion, pub, mother, just, cup, done,..}

0072 PDV //ʌ// (the majority partition for NL varieties)
states (4): ɔ̣˓ ɔ̤˓ ɑ̣. ā
e.g. cup, onion (Cf. e.g.s for OU)

0074 PDV //ω// (the majority partition for L-varieties (Northern))
states (5): ṳ˓ ọ˓ o̤˓ o̤˓ ṳ..˓
e.g. pub, cup, onion

0076 PDV //ɒ//
states (5): ɒ̣˒ ɒ̤˒ ɒ̣˓ ɒ̤˓ ɒ
e.g. hurry, worry, onion

0078 PDV //ɔ//
states (5): ɔ ɔ̤˃ ɔ̣ ɔ̤ ɔ̣
e.g. pub

0080 PDV //ʌ//
states (5): ɨ i̠ ị i̱ i̠̤
e.g. mother, just

0082 PDV //ə// (the majority partition for intermediate L/NL varieties)
states (5): ə ɯ̈ ɤ̈ ʌ̄ ə̤
e.g. cup, onion, just, done

OU {ɷ} NL {pull, put, woman, good, butcher, book, . . . }

0084 PDV //ɷ// (the majority partition)
states (4): ɷ̣ ɷ˂ ɷ̣˂ ṳ˂
Cf. e.g.s for the OU

0086 PDV //u//
states (3): ṳ ụ˂ ṳ
e.g. book, cook

0088 PDV //ə//
states (5): ɯ ə ɤ̈ ʌ̄ ə̤
e.g. good, butcher

OU {u} NL {moon, two, revolution, suit, tissue, do, you, who, school, tune, curiosity, . . .}

The composition of this lexical set is complicated by the phonological interaction of '/u/' and '/j/' – an interesting problem which has not attracted the sociolinguistic data it needs for an adequate solution. This complexity produces one of the few cases - in the 'segmental' subspace - where the classificatory independence of the PDVs does not hold. (Cf. criteria 2880 or 288 below depending on the nature of the interaction between '/u/' and '/j/'.)

0090 PDV //u//
states (6): ụ ụ‥˒ ụ˒ y̠ ỵ͜ụ ə₂͜ʉ
e.g. moon, two, beautiful, suit

0092 PDV //i//
states (4): i^(ɕ/ç)ə i̠ ^əi̠ ^əi̲
e.g. do, you, who

0094 PDV //ɷ//
states (3): ɷ̣ɑ̣. ɷ̣e̠ɨ ^əɷ (ɷ=ụ‥˒)
e.g. boot, school [partition not much bigger than this]

0102 PDV //ɩə//
states (5): ɨ i̠ i̤ i̲ i̲̤ (± ə/ɩ)
e.g. tune, curiosity

OU {eɩ} NL {eight, great, take, make, shape, railway, halfpenny, brains, straight, . . .}

0104 PDV //eɩ // (the majority partition for NL varieties)
states (4): e̞ɩ e̠.ɩ ɛ̠.ɩ ɛ̈ɩ
e.g. eight, great, take

0106 PDV //ɛ//
states (4): ɛ̠˒ e̞ ɛ̤ ɛ̣
e.g. take, make
[partition not much bigger than this]

0108 PDV //eɩə//
states (2): ɛi̠ə ɛ iə
e.g. shape, railway

0110 PDV //a//
states (5): a̤ a̠ a̠. ɛ̞ ɛ̠
e.g. halfpenny, take

0112 PDV //i// (an exclusive majority partition for L varieties)
states (4): i(ɛ) i̠ i̠ y
e.g. great, brains (see below)

1120 PDV //iə// (an exclusive majority partition for L varieties)
states (4): i̠ə e̞ə i̠ä i̠ä
e.g. great, brains (see below)

Localised speakers seem to realise the majority of the lexical set either with PDV //i// or with PDV //iə//, hence the use of indefinite 'exclusive'. By contrast PDVs 118 and 120 below seem to exhibit a different relationship.

I'd like to emphasise again that if samples of speakers showed that these perceived regularities were chimerical, the classifiability of those speakers would not be compromised. This is because the notion of 'a partition' (and its size and commonness) is not one which is built into the classificatory method; it is introduced solely to give the reader who happens not to know Tyneside some crude idea of the lexical patterning of the phonologies to be found there (A,C4).

0114 PDV //ɛɩ//
states (2):
e.g. eight, great, straight

1140 PDV //ə//
states (2): ə ɩ
e.g. Mon<u>day</u>, Thurs<u>day</u>, holi<u>day</u>

OU {əɷ} NL {phone, stone, so, smoke, nose, old, home, yel<u>low</u>, . . .}

0116 PDV //əɷ//
states (5): əɷ˂ aɤ̈ ɛɷ o̤˂ɷ əɷ
Cf. e.g.s for the OU

1160 PDV //əɩ//
states (3): əɩ ɩi i̠ː
e.g. so, no

0118 PDV //ɔː// (joint majority partition for L varieties)
states (4): o̤ ɔ̤˂ ɔ̤˂ ɔ
e.g. phone, so, smoke

0120 PDV //uː// (joint majority partition for L varieties)
states (5): ṳ(ə) ṳ˂ ṳ o̤˂(ə) o̤˂
e.g. phone, go, nose

0122 PDV //ɑː//
states (4): a̠. ɑ̤ ɑ.. ɑ̤
e.g. old, know, no, cold

0124 PDV //ɩə//
states (6): je̤ jɛ̠ jɛ̣ɩ ɩə i̠ə i̤ə
e.g. don't, stone, home

0126 PDV //ɛω//
states (4): ɛ̠ω ɛ̠ω$^{(ə)}$ a̠̣ω a̠̣ω$^{(ə)}$
e.g. bolt, cold, hope

1260 PDV //ə//
states (2): ə ɩ
e.g. pill<u>ow</u>, yell<u>ow</u>

OU {aɩ} NL {side, five, I, right, might, nine, . . .}

0128 PDV //aɩ// (the majority partition for NL varieties)
states (5): a̠ɛ̠ ɑ̣i a̤ɩ a̠ɩ aɩ
e.g. I, side, <u>chi</u>na

0130 PDV //aː//
states (4): ɑ ɑ̤ ɑ̣ ɑ̤̣
e.g. I, five

0132 PDV //iː//
states (4): i$^{(ɕ)}$ i̠ i̠. ɩ
e.g. right, night, blind

0134 PDV //eɩ// (majority partition for L varieties)
states (3): ɛ̠i̠̬ e̠̬i e̠i̠
e.g. knife, mine, side

OU {aɩə} NL {fire, tyre, trial, rel<u>ia</u>ble, . . .}

0136 PDV //aɩə// (what the majority partition is for NL varieties (0136, 0138, 0140) is unclear)
states (5): a̠ɩə a̤ɩə a̠ɩjə ɛ̠ɩjə äɩ^{h}ɐ

e.g. fire, tyre, trial

0138 PDV //ăə//
states (2): ă:ə ăə
e.g. fire, tyre, reliable

0140 PDV //a://
states (2): ă a̱.
e.g. fire, trial, reliable

1400 PDV //ɛiə// (the majority partition for L varieties)
states (4): ɛiə ɛiɛ ɛiɐ ɛiɑ
e.g. tyre, trial, reliable, fire

OU {aɷ} NL {house, now, crowd, cow, round, down, . . .}

0142 PDV //aɷ// (the majority partition for NL varieties)
states (5): ɑ̣:ɷ ɑ̣ɷ a̱ɷ a̱ ɷ ɑ̤ ɷ
e.g. house, now, crowd

0144 PDV //ɛɷ//
states (4): e̱ɷ ɛ̱. ɷ ɛ̱ɷ
e.g. house, crowd

1440 PDV //ɩɷ//
states (2): ɩɷ ɩ̣ɷ
e.g. now, cow
[partition not much larger than this]

0146 PDV //u://
states (4): ɷ ṳ˞˂ ᵊṳ˂ ṳ˂
e.g. house, mouse, round

1460 PDV //ɔɷ//
states (2): ɔɷ ʌɷ
e.g. house, loud, down

OU {aɷə} NL {flower, our, Stour, tower, hour, . . .}

0148 PDV //ɑɷə//
states (3): ɛ̠..ɷə ɑ̣.ɷə ɑɷə
e.g. flower, our

0150 PDV //ɑ://
states (4): ɑ̆ (= ä) ɑ̤.. ɑ..ə ɑ̣ə
e.g. flower, our, tower

1500 PDV //ɛɷə//
states (4): ɛɷə ɛɷɛ ɛɷɐ ɛɷɑ
e.g. flower, our

OU {ɔɩ} NL {boy, buoy, toil, noise, toy, boil, . . .}

0152 PDV //ɔɩ// (the majority partition for NL varieties)
states (6): ɔ̣ɩ ɔ̣ə ɑ̣ɩ ɔ̣ɩ ɒ̣.ᐸɩ ọ.ᐸɩ
Cf. e.g.s for the OU

1520 PDV //ɒɩ//(the majority partition for L varieties)
states (2): ɒ̤ᐸɩ ọ̤ᐸɩ
e.g. noise, toy

0154 PDV //əɩ//
states (2): əɩ ə̤ɩ
e.g. buoy, noise

0156 PDV //ɒɪə//
states (2): ɒ̣ɪə ɒ̤ɪə
e.g. boil, toil

0158 PDV //ɔːə//
states (4) ɔ̤ə ɔ̣ ɔ̣ː ɔ̣ː
e.g. boil, boy

OU {ɜ} NL {bird, fur, curl, year, girl, earth,
burner, worth, birth, . . .}

0160 PDV //ɜː// (majority partition for some NL varieties)
states (6): ɑ̣. e̠.˃ ə̣˃ ə̣˃ɽ ɑ ɜː
e.g. bird, fur, year, curl

1600 PDV //ɪə//
states (5): e̩.ə iə e̩ə i̠ ɜː
e.g. year
[partition restricted to this item ?]

0162 PDV //œ// (majority partition for some NL and some L varieties)
states (3): œ̠ ə̣˂ ə̤˂
e.g. bird, fur, curl, year

0164 PDV //ɛ(ə)//
states (3): ɛ̠.(ə) e̠(ɪ) ɛ̞
e.g. bird, girl, curl, year

0166 PDV //ɐ//
states (3): e̠. ɐ ɯ̤.˃
e.g. bird, girl, year

0168 PDV //ɔ// (majority partition for some L varieties)
states (8): ɔ̣.(ə) ọ˂ ọ ɔ̤.ᵊ ɔ̤. ɔ̣. ɵ ɵ̠
e.g. bird, birth, worth, year

0170 PDV //ɷə//
states (2): ɷə ṳ̣ə
e.g. burner, earth

0U {ɩə} NL {here, really, serious, ideal, beer, cheer, fear, . . .}

0172 PDV //ɩə// (majority partition for NL varieties)
states (5): ẹ.ə iə ẹə i̱ ɜː
e.g. here, really, serious, ideal

1720 PDV //e//
states (2): ẹ ẹ̣
serious, really

0176 PDV //iː//
states (3): iː iˑ i̤
e.g. nearly, really, serious

0178 PDV //ɒː//
states (3): (j)ɑ̣. (j)ɒ̣ ɩɒ̣
e.g. here, cheer, beer

1780 PDV //ɩɛ//
states (3): ɩɛ ɩiə iă
e.g. beer, fear, here

OU {ɛə} NL {hair, there, care, pair, wear, <u>fairy</u>, . . .}

0180 PDV //ɛə// (the majority partition for NL varieties)
states (3): e̞ə ɛ̞ə ɛ̠
Cf. e.g.s of the OU

0182 PDV //ɛ// (majority partition for some L varieties)
states (4): e̞̞ ɛ̠ ɛ̞ ɛ̠̞
e.g. care, there, pair

0184 PDV //ɜː// (majority partition for some L varieties)
states (2): ɜ̝ ɜ
e.g. there, pair, hair

OU {ωə} NL {poor, your, moor, <u>brew</u>er, <u>sew</u>er [<u>sensu</u> shit], . . .}

0186 PDV //ωə//
states (2): ω˂ ə ωə
e.g. your, poor, moor

0188 PDV //uə//
states (4): u̞ə u̞˂ ə u̝̞˂ ə u̝̞˂
e.g. poor, moor, your

0190 PDV //ɔ//
states (2): ɔ̝ə ɔ̝ː
e.g. moor, more, poor

0192 PDV //ωwɛ//
states (3): ωwɛ ω̝wɛ ω̞wə
e.g. <u>brew</u>er, <u>sew</u>er

3. Vowels: (b) Evidence from dialect writers

Some support, not always unequivocal, for the diasystemic variants and lexical subsets which are postulated above may be derived from the attempts made by writers, in terms of spelling changes, to represent the speech variety of their characters or their locality. Here I have abstracted, for each PDV which seems to be being represented, a few words from each of three sources: Sandvid (1964) (referred to as DS), Dobson (1969) (referred to as SD), and Irwin (1978) (referred to as DI). The most systematic use of variant spellings to suggest a Tyneside variety is in respect of stressed vowels, and of deleted, usually final, consonants; here I shall only show the former variants.

Clearly the letter strings of the dialect writer have a very different function from those of the phonetician. The former simply wants to suggest a certain quality of variety to the ear through the eye. Not all of the local phonological variants can be suggested, since the recognisability of particular orthographic forms (words) must not be too difficult for the reader. Presumably the writer wishes, rather, to suggest enough of the difference from, say, non-localised speech (in the form of conventional spellings), for the hearer to be able to supply the rest from his auditory imagination. I am not even suggesting that the dialect writer will necessarily be focussing attention on those variants which are, in the terms defined above, diagnostic, nor will it necessarily be the case that the representation of the particular variants are consistent in different items of even the same lexical subset. Finally, the fact that a dialect writer fails to represent a given PDV, or even a whole OU, at all in his orthographic innovations is **not** taken to be evidence for doubting that distinction, or those distinctions, which my system does embody.

The layout is the same as for the specifications above, but the states and the code numbers have been omitted for ease of reading.

OU {iː}

//iː// ---

//ɪ// ---

(OU {iː})

//ɛ// ---

//ei// ---

//ɪə/ ---

//ɪi// ---

OU {ɪ}

//ɪ// ---

//ɐ// ---

//ɪə// SD: geetars (guitars)

//ɜː// ---

//ɛə// DS: 'ees (his) [DS: ? enyuff (enough)]

OU {ɛ}

//ɛ// ---

//iː// DS: etiqueete, weel (well), insteed, deethly,
breid (bread), heid (head)
SD: weel, deed (dead), reed (red)
DI: heed, deed/deid (dead), weel, reed

//ɪ// DS: tegithor, injoy, yit (yet), ivry (every), nivvor (never),
yistida (yesterday)

(OU {ɛ})

SD: sivin (seven), yit, thim (them), aategithor, iver (ever), git (get)

DI: nivor, git, ivor

//ə// DS: onny (any), monny (many)

SD: onyway

DI: chorry (cherry)

//ɛə// [DS: ? harrin (herring), varry (very)]

OU {æ}

//æ// ---

//ɛ// DS: hes (has), hev (have), [? kin (can)]

SD: sex (sacks), hev, chep, hing (hang)

DI: hev, hing

[SD: ? byad (bad), tyab (tab)]

//a// ---

//ɔ// ---

OU {ɑ:}

//a// ---

//ɔ// DS: ore (are)

//æ// DS: laffed (laughed), eftor (after), raither (rather)

SD: efter (after), fethor (father)

DI: eftor, fethor

OU {ɒ}

//ɒ// ---

//ɔ// ---

//a// DS: cass ('cos), wrang, wabbling, watt (what), lang, wazn't
SD: av (of), lang (long)
DI: wrang (wrong), belang (belong)

//ə// DS: cannit (cannot)
DI: whit (what)

//ɛ// DS: het, fre' (from), wesher
SD: wes (was), het
DI: wesh (wash)

//ʌ// ---

//ω// DS: wuz (was)

OU {ɔ}

//a// DS: aaful, aal, scraal, straa, claa,
wattor (water), scad (scald), tallk (talk)
SD: caal (call), aaful, haak (hawk), saa (saw),
watters (waters), scaad (scald)
DI: aal, taak (talk), waalk (walk), waarm, caal, laa (law)

//ɔ// ---

//ɒ// ---

//ɛ// DS: mair (more)
SD: yer (your), fer (for)

//aωə// DS: fower (four), bowt (bought), thowt (thought)
SD: fower
DI: bowt, thowt, fowerty (forty), dowter (daughter)

//ωə// - - -

OU {ʌ}

//ʌ// - - -

//ω// DS: (? [u:]) sooks (sucks (n.)), (?[jω]) enyuff (enough)
SD: cum (come), ennuf, muthor, munny
DI: (? [u:]) dookin (ducking), muthor

//ɒ// DS: thorrily, knockle, amang (among)

//ɔ// - - -

//ɩ// DS: disn't, sitch, jist,
DS: ? dyen (done), ? nyen (none)
SD: deun (done), shyeul (shovel)

//ə// - - -

OU {ω}

//ω// - - -

//u// DS: wull (wool), luk
SD: luk, gud, byeuk (book)

DI: shud (should)

//ə// DS: wud (would), wadn't (wouldn't)
SD: wad (would)
DI: waald(would)

OU {u}

//u// ---

//i// DS: dee (do), twee (two), te (to), whe (who)
SD: te (to), de (do), ye (you), whe (who)
DI: te, ye, whe, dein/deein (doing)

//ɒ// DS: loss (lose)
SD: thro (through)

//ɪə// DS: byets (boots), fyul (fool), tyuth (tooth), syun (soon), skyul (school)
SD: fyeul (fool), byeut (boot)
DI: scyul (school), buyets (boots)

OU {eɪ}

//eɪ// ---

//ɛ// DS: the' (they), m<u>e</u>vvies (maybe), br<u>e</u>vely
SD: tek (take)

//eɪə// ---

//a// ---

//i// DS: geet (great), eethor (either), bleeze, aalwiz (always)
DI: bleezer, greet

//iə// DS: hyemly (homely), myeks (makes), tyeble (table), agyen (again), pyeppor (paper), yebble (able)
SD: spayus (space), agyen (again), tyek (take), shyem (shame), syem (same)
DI: tyeke (take), agyen (again), shyem (shame)

//ɛɪ// DS: ighty (eighty), stright (straight)
SD: stright
DI: fyce (face), myke (make)

//ə// ---

OU {əʊ}

//əʊ// ---

//əɪ// DS: maist (most), cla'es (clothes), waint (won't), se (so), dain't (don't), nee (no), swall_y (swallow), se (so)
SD: ne (no)
DI: ne (no), claes (clothes), see (so)

//ɔː// DI: blaws (blows)

//uː// ---

//ɑː// DS: knaa, blaa, aan, gaan (going), aad (old)
SD: aad, knaa, had (hold)
DI: gaa (go), raa (row)

//ɪə// DS: styens (stones), byeth (both), hyemly (homely)
SD: hyem (home)

DI: hyem

//ɛɷ// DS: rowl (roll), sowl (soul)
SD: sowljers (soldiers)
DI: towld, howld (hold)

//ə// SD: barr<u>er</u> (barrow)
DI: n<u>e</u>body

OU {aɩ}

//aɩ// ---

//ɑː// DS: aa'm (I'm), ma (my)
SD: ah (I)
DI: aa (I)

//iː// DS: reet, neet, neethor (neither), dee (do),
(?[ɩ]) ah<u>i</u>nt (behind), finnd (find), blinnd (blind)
SD: aalreet
DI: reet, neet, me (my)

//eɩ// DS: me (my), m<u>e</u>self
SD: whey (why)

OU {aɩə}

//aɩə// ---

//ãə// ---

//aː// ---

//ɛiə// - - -

OU {aɷ}

//aɷ// - - -

//ɛɷ// - - -

//ɩɷ// - - -

//uː// DS: Sooth (south), hoo, roond, oot, doon, broon,
(? [ɷ]) fund (found), pund (pound)
SD: doon, groond, aboot, hoo (how), shooting (shouting),
broon, coont (count)
DI: noo, aboot, toon, roond, hoose, anyhoo, coo,
(? [ɷ]) pund

//ɔɷ// - - -

OU {aɷə}

//ɑɷə// - - -

//ɑː// - - -

//ɛɷə// [DS: ?[ɔ] wor (our)]
[SD: ?[ɔ] wor]
[DI: ?[ɔ] wor]
DI: flooers (flowers) (=?[uːə])

OU {ɔɩ}

//ɔɩ// - - -

//ɒɩ// ---

//əɩ// ---

//ɒɩə// ---

//ɔːə// ---

OU {ɜ}

//ɜː// ---

//ɩə// ---

//œ// ---

//ɛ(ə)// ---

//ɐ// DS: s_a_rcumstances, cons_a_rn
DI: sartinly, shart

//ɔ// DS: pref_o_rences, s_o_rviettes, tegith_o_r, torn (turn), stor (stir), wawm (worm), wawk (work)
SD: jawney (journey), torn (turn), distorb (disturb), morder (murder), bord (bird), hor (her)
DI: sorr (sir), hor, forst (first), hord (heard), dorty, morder, Porcy (Percy).

//ɷə// ---

OU {ɩə}

//ɩə// ---

//e// ---

//iː// DS: reelise, yeer
[DS: ? [ɩ] inj_i_ns (indians)]

//jɔː// DS: deor (dear), queeor, heor (here), neorin (nearing),
heor (hear), dreory (dreary)
SD: beyor (beer)
DI: beor (beer), heor (here), cheor (cheer)

//ɩɛ// ---

OU {ɛə}

//ɛə// ---

//ɛ// ---

//ɜː//
[DS: (? [ɔ]) thor (their), dore (dare)
(? [ɩɷ]) teor (tear (v.)) wheor (where)]
[SD: (? [ɔ]) thor, thurselves (theirselves)
(? [ɩɷ]) theor (there), wheor]
[DI: (? [ɔ]) thor]

OU {ɷə}

//ɷə// ---

//uə// ---

//ɔ// ---

//ɷwɛ//- - -

Non-reduced unstressed vowels:

{ə}

ɒ DS: propor, pyeppor (paper), connoshes (connoisseurs)
SD: muthor
DI: mistor, pictor (picture), lavotary

ɩ DS: despritly
SD: massacree, deppity (deputy)
[? [ɛ] inter (into)]

Here I had also intended to discuss at some length the extent to which Ellis' (1889) account of what he called the **North Northern** dialects gives interesting and useful time-depth support to the criteria proposed above. (I believe that the support, though not wholly definite, is at both the PDV and the State levels.) There are, however, considerable, but not insurmountable, problems of interpretation attaching to 'palaeotype', the name which Ellis gave to his transcription system (and its symbols). Local (1981), however, points out that many of the problems associated with palaeotype arise from trying to translate it into a simple unitary framework (such as the IPA). By examining the evolution of palaeotype, Local shows that it is "a complex mixture of the phonetic and the phonological, the systematic and the fortuitous" (1981: 2), and he claims that in showing this, he will "demonstrate that it is no more 'incomprehensible' or 'defiant of interpretation' than any other attempt to represent sounds by symbols on paper" (1981:2). This may well be the case, but because of the mixture problem, the interpretation of any given chunk of transcription often may not be unary. Needless to say, though Local (1981) establishes his general case, he can give nothing like a full catalogue of those segments of any transcription which are phonetic, those which are phonological, or those which are fortuitous, and we would surely be surprised were it the case that a given symbol in palaeotype always had only one of these qualities

The indeterminacy of Ellis' system, and the consequent range of possible interpretations which can be derived from it, is such that I have decided to forego the pleasures of a lengthy analysis. Nevertheless, for the sake of incompleteness, and for the reader's delectation, I have reproduced, in Figure A1 (5 pp.) and in Figure A2 (7 pp.), respectively, his comparative transcriptions of one text from three locations (South Shields, Newcastle upon Tyne, Berwick upon Tweed), and of a different text from twenty two locations (of which Newcastle is, again, one). Flavour, if not uniquely definable substance, can certainly be got from an imaginative reading of these transcriptions. Trying to assign form, and then meaning, to the differences between the transcriptions is what causes the problems.

4. Criteria for centralised vowels in unstressed syllables

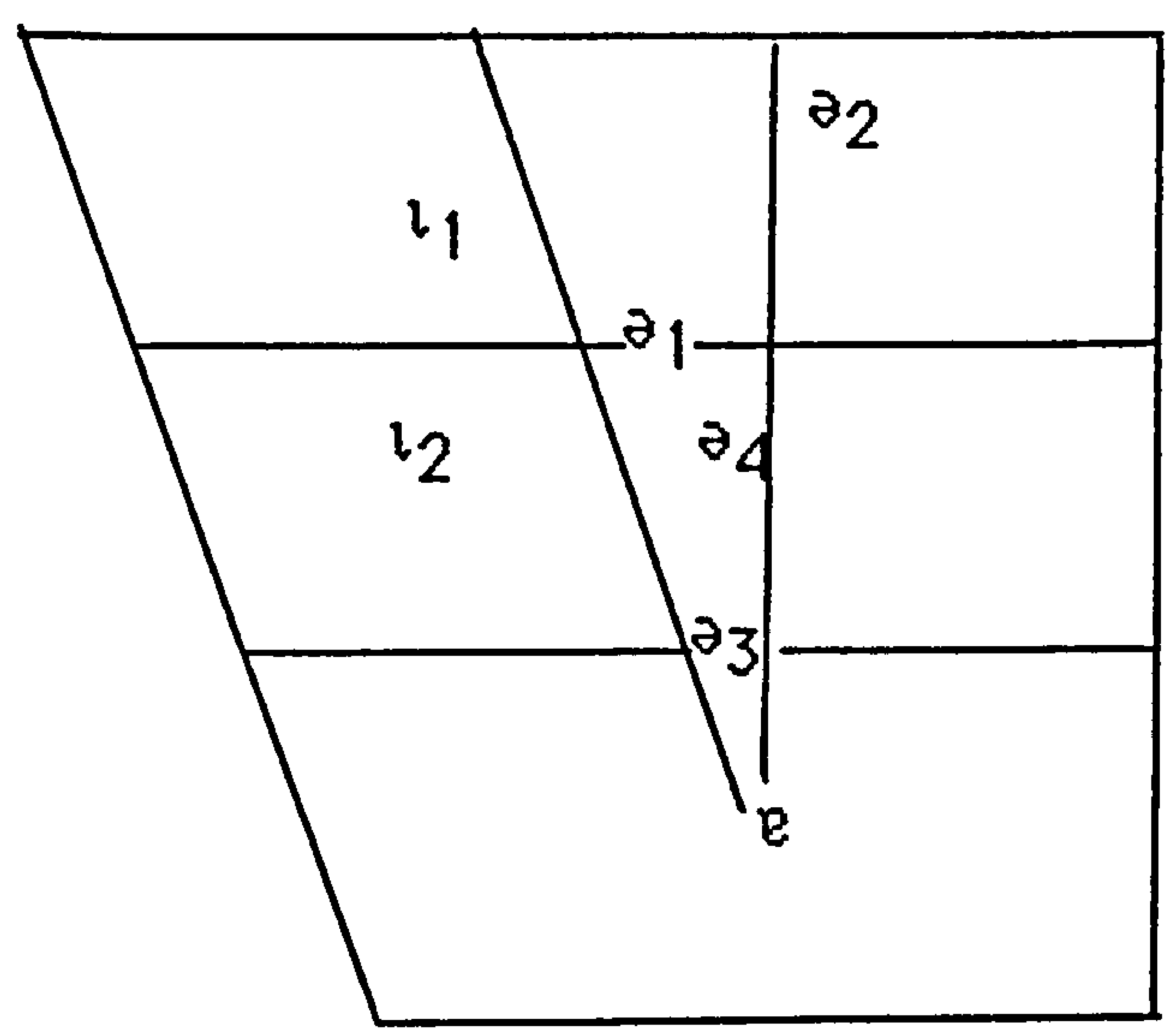
These criteria cover the vowels of unstressed syllables. The basis for the definition of the OUs is made with reference to RP, but there are in addition some purely phonological sub-classifications. Thus [$ə_3$], using Jones's notation (1976: 91 ff.), and [$ə_{4a}$] are distinguished not on the grounds that they have distinct realisations in RP - which may or may not now be the case - but because the phonological contexts are distinct, and because the former can be realised with [ɐ]. Nevertheless I do not preclude the possibility of a strong connection between them in terms of underlying rules, e.g. choice of realisation under criterion 196 might strongly predict a similar choice under criterion 200. In line with other contexts of the same principle - as the reader has no doubt guessed - I propose to keep them separate on the grounds that it is precisely these assumptions of co-occurrence which need to be documented rather than left in a state of suspended, even if plausible, assumption (A,C5).

Reasons for not following an RP basis of the OU principle for these criteria are not far to seek, notably both the extreme unreliability of one's perception of the phonetic difference between centralised vowels, and a general doubt that these differences are in any sense the real basis even for Jones's <u>own</u> classification; rather it seems that the phonological contexts - to which the lists of exemplars are pointers - form the basis for the different phones which he introduces.

Here these contexts have merely been made explicit, though in a fairly non-committal form - e.g. [$ɩ_1$] and [$ɩ_2$] follow the RP scheme more closely than [$ə_3$] and [$ə_{4a}$] - the most important guiding principle having been that information concerning the relationships between the criteria be recoverable from the analysed data. The scheme presented here thus tries to avoid representing information about what might turn out to be distinct processes in one and the same criterion.

PDVs are defined for these criteria in terms of the dichotomy <u>reduced /unreduced</u> since it seems to represent the major distinctions between speakers. In a pilot study of these phenomena McNeany (1972) found that whilst intra-speaker mixing of what are here represented as states was possible, intra-speaker mixing of reduced and unreduced forms was rather rare.

The relative positions of the various unstressed neutral vowel states (Jones 1976: 67 ff. & 91 ff.) are indicated by:



McNeany (1972) confronts the theory of Chomsky and Halle (1968) with data from the realisational variation in vocalic nuclei of unstressed syllables. In doing so, he proposes certain changes to the theory. Whether or not a particular speaker reduces the vowel in an unstressed syllable (in speech of a given speed) is a matter which varies systematically from one variety to another. It is therefore a matter of competence rather than performance.

Chomsky & Halle assume that it is reasonable and sufficient to assign exact feature-specifications to a non-terminal category such as reduced, but McNeany shows how this assumption is very seriously weakened by his interpretation of his data.

McNeany (1972) points out that the three forms [ɹiˈzɒlv] / [ɹɩˈzɒlv] / [ɹəˈzɒlv], and many other similar sets of items, are all current in $_L$T varieties. He sets up a category [± reduced] to be incorporated in the phonological rules of the grammar, specifying three degrees of this category, namely [1 reduced] (for [ə]), [2 reduced] (for [ɩ]), and [unreduced]. In his collection of data, McNeany noticed that even in the uttering of word-lists there is very little conscious correction of the patterns of reduction/non-reduction which are found in naturally occurring speech (1972: 9).

The main difficulty, for our present purpose, is how one can establish the lexical sets in which the processes of reduction/non-reduction occur in different ways. Looking at the patterns of reduction/non-reduction in many naturally spoken different lexical items, McNeany tried to group those items in terms of phonological environments. He then established empirically quantitative expectations in respect of the likelihood of reduction/non-reduction for these environments in L-varieties and NL-varieties. Such expectations were only an expression of general tendency. The extent to which those rules were obeyed could then be taken as a variety-determining, or a variety-differentiating, factor.

In his initial work (1972), McNeany specified five lexical sets of items (I,II,III,IV,V). Each lexical set was variously subdivided. He exemplified them as follows (McNeany 1972: 10-11), suggesting that though the examples are typical, they should not be thought of as representing "anything more than one possibility amongst many":

Lexical sets:

			Localised	Non-localised
I	(a)	paper	[ˈpeᵊpɛ]	[ˈpeɪpə]
		number	[ˈnɤmbɛ]	[ˈnʌmbə]
		ever	[ˈɛvɛ]	[ˈɛvə]

	sofa	['sofɛ]	['səɷfə]
	Pamela	['pämlɛ]	['pamlə]
(b)	interview	['ɪntɛvjɪɛu]	['ɪntəvjɷu]
	afterwards	['äftɛwɛdz]	['aːftəwədz]
(c)	method	['mɛθɛd]	['mɛθəd]
	coward	['kɛwɛd]	['kaɷəd]
	concert	['kɒnsɛt]	['kɒnsət]
II (a)	extremity	[ɛks'tɹɛmətɪ]	[ɪks'tɹɛmɪtɪ]
	irregular	[ə'ɹɛgl̩ɛ]	[ɪ'ɹɛgjɷlə]
	solid	['sɒləd]	['sɒlɪd]
(b)	problem	['pɹɒbləm]/['pɹɒblɪm]	['pɹɒblɪm]
	system	['sɪstəm]	['sɪstim]
	curtain	['køtn̩]	['kɜːtɪn]
	houses	['hɛɷzəz]/['hɛɷzɪz]	['haɷzɪz]
(c)	impression	[ɪm'pɹɛʃn̩]/[ɪm'pɹɛʃɪn]	[ɪm'pɹɛʃn̩]
	seven	['sɛvɪn]	['sɛvn̩]
	almond	['aːmɪnd]	['aːmn̩d]
	equipment	[i'kwɪpmɪnt]	[ɪ'kwɪpmənt]
III (a)	party	['paːti]	['paːtɪ]
	city	['sɪti]	['sɪtɪ]
(b)	elaborate	[ɪ'labəɹət]	[ɪ'labəɹɪt]
	defend	[di'fɛnd]	[dɪ'fɛnd]
IV (a)	perfect	['pøfɛkt]	['pɜːfɪkt]
	subject	['sɤbdʒɛkt]	['sʌbdʒɪkt]
(b)	explain	[ɛks'pleᵊn]	[ɪks'pleɪn]
	expose	[ɛks'poz]	[ɪks'pəɷz]
	engage	[ɛn'geᵊdʒ]	[ɪn'geɪdʒ]
V (a)	contend	[kɒn'tɛnd]	[kən'tɛnd]
	control	[kɒn'tɹol]	[kən'tɹəɷl]
	contemplative	[kɒn'tɛmplətəv]	[kən'tɛmplətɪv]
(b)	compensation	[kɒmpɛn'seᵊʃn̩]	[kɒmpən'seɪʃn̩]
	accent	['äksɛnt]	['aksənt]

(c)	totality	[to'tälәti]	[tә'talɪtɪ]
	vocation	[vo'keәʃn̩]	[vә'keɪʃn̩]
(d)	ambassador	[ɑm'bɑsәdɛ]	[әm'basәdә]
	transcription	[tɹäns'kɹɪpʃn̩]	[tɹәns'kɹɪpʃn̩]
(e)	tomorrow	[tә'mɒɹә]	[tә'mɒɹәʊ]
	yellow	['jɛlә]/['jɛlɪ]	['jɛlәʊ]

The following criteria (428-431) examine the characteristic patterning (for each speaker) of vowel reduction / non-reduction in lexical sets of weak forms of largely closed system items which, during the historical development of the Tyneside Linguistic Survey, were not originally included at this point in the list of criteria. (Hence the apparently erroneous numbering.) The original location for them was in what is here Appendix C, below.
Each lexical set is presented as a matrix whose properties are discussed below it, as follows:

Weak forms of:
from /for /upon /or /your /my /he /someone /go /be /been /are /his

1. DT / / / / / / .....
(unreduced)

2. DU / / / / / / .....
(reduced)

Majority
(coded 1 or 0 or NC)
3. DT(0)}
4. DU(1)} / / / / / .....
5. NC }

Each token of each weak form is coded as being reduced (row 2) or unreduced (row 1), since McNeany (1972) showed that the strong (stressed) or weak (unstressed) nature of the item is perceived independently of the reduced or unreduced nature of the vowel. Then these frequencies are summarised for the type of any given weak form for the given speaker as being either, by majority, reduced (DU, represented as (1)), or unreduced (DT, represented as (0)).

428 ((Total no. of (1)s in row 4)/(15 - (NC)))100

The criterion expresses the percentage of items in the list having a majority representation of reduced (neutral) vowels when they are weak (unstressed).
[The maximum in the denominator is artificially high (i.e. 15 for 13 items) to accommodate items that may need to be included later. This means that the percentage won't reach 100 unless a dummy NC element is used.]
This criterion measures the speaker's range of reduced forms in terms of the number of items in the list which are typically reduced, rather than measuring the exact frequency of reduction of each item in the set. Clearly, both are relevant measures for a sociolinguistic survey. But because the latter will tend to measure variety mixing of a certain kind, and the former will tend to express a varietal norm (because it throws away much of the presumed continuity of the distribution), we prefer to consider only the former, at least initially (A,C6). The form of the criterion expressed here is, in the present state of the art, both more discriminating and better suited to the retrieval of the relevant information (since the different items are kept separate throughout the coding process). There is evidence in McNeany (1972) that Tyneside L-variety speakers will score low on 428 (i.e. have few reduced forms).

Weak forms of:

/but/just/some/saint/had/have/shall/should/could/would/were/was

1. DV
(unreduced)
/ / / / / / .....

2. DX
(reduced)
/ / / / / / .....

Majority
(coded 1 or 0 or NC)
3. DX(1) }
4. DV(0) } / / / / / .....
5. NC }

429 ((Total no. of (1)s in row 3)/(14 - (NC)))100

There is much less firm evidence about how this set behaves, though it is fairly certain that there are varietal differences of some significance. The contexts in which alternation is possible may be limited in ways which are not yet fully understood. LT-varieties will tend to use unreduced forms, but this is less cetain than for the 428 set. Peculiar reversals of this measure may be uncovered which will doubtless be interesting.

Weak forms of:
we/she/they/one/till/Gateshead/Newcastle/you/so

1. DV
(unreduced) / / / / / / .....

2. DZ
(reduced) / / / / / /.....

Majority
3. DZ(1) }
4. DY(0) } / / / / / /.....
5. NC }

430 ((Total no. of (1)s in row 3)/(11 - (NC))100
The prediction, from McNeany (1972), is that these items are almost universally reduced in certain $_L$T-varieties.

EA Total no. of weak realisations of:
/as/at/of/to/so/such/do/does/was/could/should/would/can
/you/we/me (realised as [əz])
EB Total of above realised with [ɨ̞]
431 (EB/EA)100

In this criterion the notion of weakness is somewhat relaxed since it is far less clear that one is here dealing with a reduction/non-reduction phenomenon. The [ɨ̞] realisations we are interested in here are not, in fact, confined to unstressed positions, cf. "when I was 'at [ɨ̞t] school". Nevertheless, though this is the case, it would clearly introduce a good deal of redundancy if we attempted to incorporate such realisations in the OU (PDV) system, especially since the variant appears to be restricted to this item alone.

OU $ə_3$ final open {bak<u>er</u>, chin<u>a</u>, trait<u>or</u>, . . .}
(But cf. $ə_2$, $ə_1$ / [non-fortis] ____*, below.)

0194 PDV //reduced//
states (2): $ə_3$ $ə_4$
e.g. (NL) baker, china

0196 PDV //non-reduced//
states (4): ɛ̱ ɩ ɐ ɑ
e.g. bak<u>er</u> (1,3,4), chin<u>a</u> (1,2,3,4), Sandra (3,4)

OU [$ə_{4a}$] /____C*
/[fortis]______(r)*C_0V . .   (But cf.[$ə_{4b}$] below)

0198 PDV //reduced//
states (2) $ə_3$ $ə_4$
e.g. stand<u>ard</u>, int<u>er</u>view (1,2)

0200 PDV //non-reduced//
states (4): $ə_3$ $ə_4$ ɛ̱ ɩ
e.g. stand<u>ard</u> (1,3), int<u>er</u>view (1,2,3,4)

OU [$ə_2$, $ə_1$] /[non-fortis]_____*

0202 PDV //reduced//
states (2): $ə_1$ $ə_2$
hamm<u>ock</u> (1,2), pave<u>ment</u>, acc<u>ent</u>, alm<u>ond</u> (1)

0204 PDV //non-reduced//
states (2): $ɩ_1$ $ɩ_2$

e.g. hamm<u>ock</u>, pave<u>ment</u>, alm<u>ond</u> (1,2)

0206 PDV //non-reduced//
states (1): ε
e.g. acc<u>ent</u>

OU [$ɪ_1$] /____C

0208 PDV //reduced//
states (3): $ə_1$ $ə_4$ $ɪ_2$
e.g. hous<u>es</u>, plac<u>es</u>, stup<u>id</u>

0210 PDV //non-reduced ɪ// (Note *)
states (2): ε̤ iˑ
e.g. <u>ex</u>pect, per<u>fect</u> (adj.)(1), <u>de</u>mand, <u>er</u>ect (2)

OU [$ɪ_2$] /____**

0212 PDV (Note *)
states (5): ɪ ɪi i əɪ ε(ɪ)
e.g. part<u>y</u>, city (1,2,3,4,5)

---

* Note. In the case of both of these criteria (0210, 0212) we must allow multiple coding on single occurrences of lexical items. Of course this may be castigated as an inconsistency in the coding system. A more positive view would see that where the definiteness of the variable is itself in doubt it is appropriate that the coding system become a little flakey (A,C7).

OU [$ə_{4b}$] / [+low][+tense] CC [1 str]

2000 PDV //reduced//
state (1): ə
e.g. observe, object (vb.)

2002 PDV //non-reduced//
state (1): ɒ
e.g. observe, object (vb.)

[In spite of Daniel Jones's inclusion of [ŭi] as a weak diphthong (1976: 85 & 125) to account for such items as fluid, ruin, Bruin, etc., it has been excluded as an OU on two grounds:
(1) relevant sociolinguistic discriminations of subsets of this lexical set appear not to differ from those which would be captured by taking the diphthong to be represented by both OU {u} and its PDVs (0090 - 0102), and OU [$ɩ_1$ /___C] and its PDVs (0208, 0210).
(Jones himself suggests the reasonableness of such a view in the form of a method for foreign learners (1976: 125)).
(2) There seem to be good reasons (Hockett 1955) for regarding this as two peaks, not one.]

5. Criteria for consonants

These fall into four groups. The first group reflects the importance of word-positional factors to the sociolinguistic discrimination of subsets of the lexical set (0214 - 0248). The second group reflects the importance of phonetic environment to the sociolinguistic discrimination of subsets of the lexical set (0278 - 0284, 0288, 2880). The third (small) group reflects the impact of morpheme type on the sociolinguistic discrimination of subsets of the lexical set (0276, 2760). The fourth group of consonantal criteria is undifferentiated as to intervening variables (due to their lack of variability, or my ignorance of it, or both) (0250 - 0274, 0286, 0294).

The first, positional, group of criteria employs senses of 'initial', 'medial' and 'final' which require some elucidation. 'Initial' means prevocalic (whether as part of a cluster or not). 'Medial' signifies either alone and intervocalic, or postvocalic before a syllabic consonant. 'Final' means postvocalic (whether as part of a cluster or not). The first and the last distinctions apply to syllables, 'medial' applies in the absence of free-morpheme boundary.

Coding conventions

The treatment of the first group of consonantal criteria (positional) is somewhat different from that of vocalic PDVs, since a numerical threshold is imposed. Any state of a positional PDV which occurs in 25% of the occurrences of that PDV is coded. The figure of 25% is of course arbitrary and must therefore be subject to adjustment in the light of later experience. For the other groups of consonantal criteria (phonetic environment, morpheme type, undifferentiated) one occurrence of a state is taken to be sufficient basis for its coding. This is rendered necessary in part by the large majority of realisations falling in state (1) - cf. 0250(1), 0252(1), 0254(1), 0256(1), etc., etc.

Notation

Examples are often followed by parenthesised numbers which refer to the most characteristic states (realisations) for that particular lexical item and its associated subset (if any) of the lexical set.

OU p ^{NL} {pot, spy, happy, capital, dip, . . .}

0214 PDV //p initial//
states (6): p^{hh} p^{h} p b̥ ɓ p'
e.g. pot (1,2,3), spy (3,4,5,6)

0216 PDV //p medial//
states (6): b b̥ p^{h} p p(ingr.) ± ɸ/β
happy, capital (1,2,3,4,5,6)

Some medial realisations occur with homorganic friction (h.a.). State (6) can therefore be coded together with other states - e.g. happy with 0216(2+6) [hăb̥‿ɸɩ]. Some co-occurrences do not apply. This general principle - ± h.a. - applies also to the rest of the stop series.

0218 PDV //p final//
states (7): b̥ ʔ‿p p p' p^{h} ᶲp p(ingr.)
e.g. dip (1,2,3,4,5,6,7)
(4 is ejective; 6 has audible closure; 7 is ingressive as in medial position. In examples of this last, for example top up where the first p is ingressive, we may be seeing the process of low level re-analysis (or equivalently the fusing of units). That is, that the item should possibly be analysed under 0216 rather

than 0218. Until further information is forthcoming however, we shall let the PDVs dominate their sets without reallocation of what we can only assume <u>may</u> be special cases.

OU b^{NL} {bag, bin, bring, robber, ribbon, dab, rub, . . .}

0220 PDV //b initial//
states (4): b b̥^{h} p(b̥̥) ± h.a.
e.g. bag, bin, bring (1,2,3,4)

0222 PDV //b medial//
states (4): b b̥ $^{\beta}$b ± h.a.
e.g. robber, ribbon,

0224 PDV //b final//
states (4): b(ə) b̥ $^{\beta}$b ± h.a.
e.g. dab, rub (1,2,3,4)

OU t^{NL} {toss, stint, letter, matter, button, hat, sit, . . .}

0226 PDV //t initial//
states (4): t d̥ t^{h} ± h.a.

0228 PDV //t medial//
states (9): ʔ t t^{h} d d̥ ɾ (t)θ ∅ ± h.a.

e.g. letter, matter (1,2,3,4,5,6,7,8,9)

0230 PDV //t final//

states (8): t tʰ t̰ t' d ɾ̥ ʔ ± h.a.

hat, sit (1,2,3,4,6,7,8)

Prince Consort Road (5)

OU d {dish, did, window, rudder, hidden, lad, . . .}

0232 PDV //d initial//

states (3): d̥ d(d̥) ± h.a.

e.g. dish, did (1,2,3)

0234 PDV //d medial//

states (6): d d̥ t ʔ ɾ ± h.a.

e.g. rudder (1,2,6), window (4), hidden (3,6), ladder (5)

0236 PDV //d final//

states (6): də d̰ d d̥ t⁽ʰ⁾ ± h.a.

e.g. lad, red (1,2,3,4,6), stupid, husband (5,6)

OU k {kit, cope, skin, kicker, Byker, sack, buckle, . . .}

0238 PDV //k initial//

states (6): k cˤ kʰ k̠ g̊ ± h.a.

e.g. kit, cope (1,2,3,4,6), skin (1,4,5,6)

0240 PDV //k medial//

states (6): k kʰ ˣk ʔ͜k Ø ± h.a.

e.g. kicker, Byker (1,2,3,4,5,6)

(3) = audible closure (onset friction);

cf. also 0242(2) & 0246(3).

0242 PDV //k final//

states (10): k ^{x}k χ g̊ q k'

± h.a. Ø ʔ‿k ʔ

e.g. sack (1,2,3,4,5,6,7,10), ask (8)

[for state (9) cf. note on 0218(7) above]

OU g {ground, gape, bigger, boggle, log, fig, . . .}

0244 PDV //g initial//

states (4): g g̊ ɣ ± h.a.

e.g. ground, gape (1,2,3,4)

0246 PDV //g medial//

states (5): g ḡ g^{x} ʔ ± h.a.

e.g. bigger, boggle (1,2,3,4,5)

0248 PDV //g final//

states (3): g^{ə} g̊ g

e.g. log, fig (1,2,3)

OU tʃ {church, change, itch, catch, French, . . .}

0250 'PDV'[N] tʃ

'states' (5): tʃ t̤ʃ t^{s}ʃ t^{s} ʃ

e.g. church, change (1,2,3,4), French (5)

[N] In this OU and some of the following, the relationship between the PDV and the OU is a fixed one-one relation. That is, the single 'PDV' has the whole set as its 'partition' of the lexical set. Informants should clearly not be scored as sharing a similiarity on the basis of sharing such a 'PDV', since, given that they are comparable, their similarity would be spuriously increased by such treatment. This kind of variable is therefore expressed by writing 'PDV' and 'state' rather than PDV and state. One occurrence of any state is sufficient for its coding. This method of representation is necessary because of our socio-perceptual ignorance of the distributional consistency of variants of these criteria (A,C8). 7

OU dʒ {jury, judge, fudge, . . .}

0252 'PDV' dʒ
'states' (3): dʒ d̥ʒ dᶻʒ
e.g. jury (1,2,3)

OU f {fetch, fresh, half, wharf, . . .}
0254 'PDV' f
'states' (5): f f' f̫ v ɸ
e.g. fetch (1,2,3), half, wharf (4), fresh (5)

OU v {very, drove, . . .}
0256 'PDV' v
'states' (3): v f w (= υ)
e.g. very (1,3), drove (2)

OU θ {thing, Arthur, lath, anything, wreath, wrath, . . .}
0258 'PDV' θ
'states' (6): θ θ̣ ··· tθ t̪ ð̃ n̪
e.g. thing (1,2,3), Arthur, lath (4), anything (5,6)

OU ð {them, then, with, loathe, . . .}

0260 'PDV' ð

'states' (6): ð θ ᵈð v n̤ ð̃

e.g. them (1,3,4), with (2), then, (5,6)

For 'states' (5) & (6) it is important to distinguish between the contexts of candidates for this and for another criterion. A realisation of (5,6) in the context in them belongs in criterion 368 (progressive assimilation) in Appendix C below. Here we are examining contexts such as some of them, where there is either no close, or no apparent, environmental motivation for the nasalisation of the dental fricative.

OU s {soup, business, listen, . . .}

0262 'PDV' s

'states' (4): s s^{s} s̬ z

soup (1,2,3), business, listen (4)

OU z {zinc, zany, rose, hose, . . .}

0264 'PDV' z

'states' (2): z s (= $s̬^{s}$)

e.g. rose, hose (1,2)

OU ʃ {ship, rush, . . .}

0266 'PDV' ʃ

'states' (2) ʃ (ʃ̥) $sʃ^{(s)}$

e.g. ship, rush (1,2)

OU ʒ {garage, television, pleasure, . . .}

0268 'PDV' ʒ
'states' (3): ʒ ʃʒ dʒ
e.g. garage, pleasure (1), television (2), garage (3)

OU h {happy, heavy, . . .}

0270 'PDV' h
'states' (4): h ɦ hˤ ø
e.g. happy, heavy (1,2,3,4)

OU m {Mary, hammer, human, home, buxom, . . .}

0272 'PDV' m
'states' (3): m m̆ ɱ

OU n {nice , rain, unpaid, unkind, sunburn, mine, main, . . .}

0274 'PDV' n
'states' (7): n nː n̆ m n̪ ɱ ŋ
e.g. nice, rain (1,2,3,5), unpaid, sunburn (4,6),
unkind (7)

OU ŋ (in free morpheme)

0276 'PDV' ŋ

'states' (7): ŋ ŋg ŋ̆ (= ŋ̆ᵏ or ŋ̆ˀ) n ng ŋk k

e.g. sing (1,2,4), singer (3,5), something (6,7)

OU ŋ (in bound morpheme)

2760 'PDV' ŋ

'states' (3): ŋ ŋg n

e.g. feeling, singing (1,2,3)

(For this PDV we revert to 25% of realisations as the threshold for coding the presence of a state.)

OU l {like, filling, cloud acclimatise, old, cold, bottle, . . .}

0278 PDV /{V}____ V . . . . .
{j }

states (7): l l̠ l̲ l̯ ɭ ɮ Ø

e.g. like (1,2,3,4,6,), filling (5,7)

0280 PDV /( ) C_0 ______V . . . . .

states (6): l l̠ l̲ l̯ ɭ ʔ

e.g. cloud (1,2,3,4,5), acclimatise (6)

0282 PDV /V______ {#}
{C}

states (5): ɭ ɮ ɬ Ø l

e.g. old, cold (1,2,3,4,5)

0284 PDV /( )____*
states (5): ɬ ɭ l ɮ Ø
e.g. bottle (1,2,3,4,5)

OU r

0286 'PDV' r
'states' (11): ɹ ɾ χ ʏ r̊ r ɽ Ø ɻ ʃ ʋ
e.g. rich (1,6,11), Harry (2,3,4,6,7,11),
tribe (5,6,8,9,10,11)

OU j

2880 PDV /alveolar { stop }________ u . . .
{fricative}

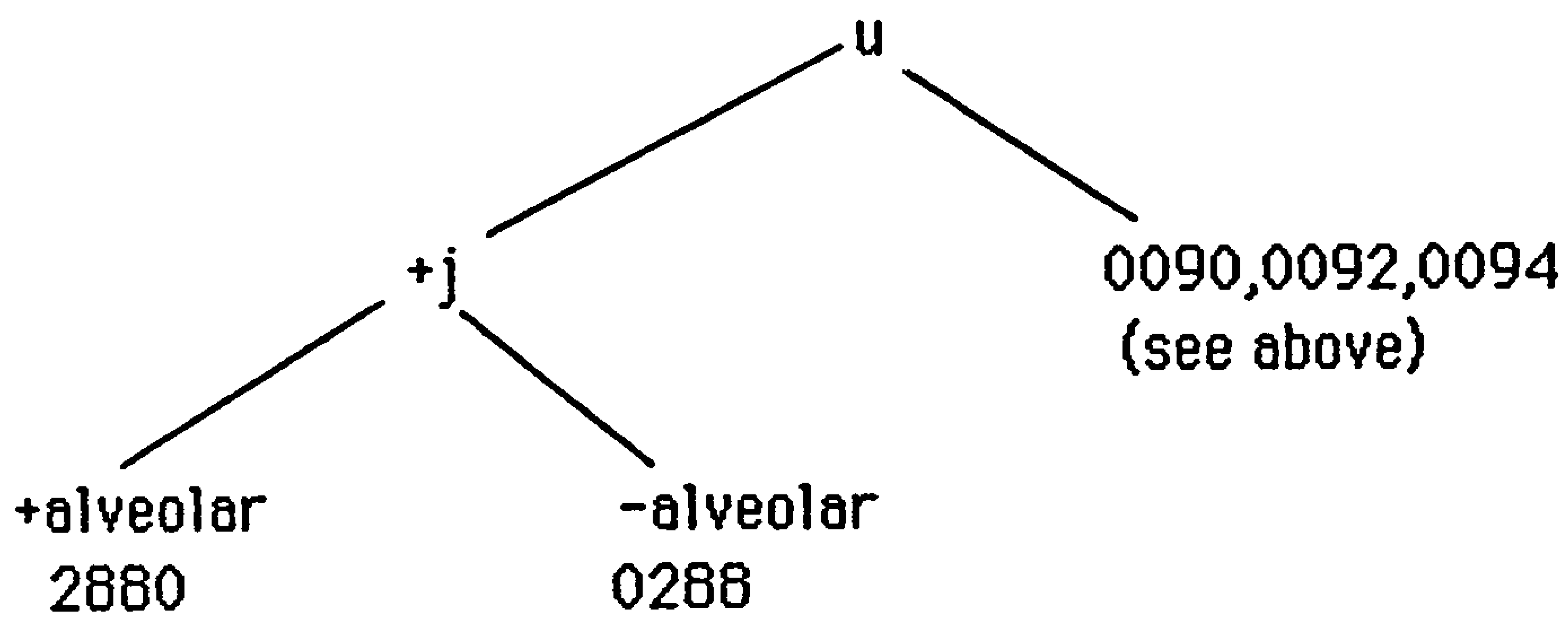
states (4): Ø j F j*

where F = [± voice] [+alveolar palatal] [+fric]
and j* = F + j;
e.g. tulip, dew (1), tulip, dew, issue, assure (2,3),
tulip, issue (4):

[tu:lɪp] [du:]
[tju:lɪp] [dju:] [ɪsju:] [əsjɷə]
[tʃu:lɪp] [dʒu:] [ɪʃu:] [əʃɷə]
[tʃju:lɪp] [ɪʃju:]

This PDV includes all instances of [j] (preceded by alveolar stop or fricative and followed by [u]) which have a potential for affrication or deletion.

The codes of these and related environments are:



0288 PDV //contexts other than above//
states (3): j $j^{ʕ}$ Ø
e.g. pure, cure, furious (1,2)
enthusiasm, revolution (3)

OU w

0294 'PDV' w
'states' (8): w hw ʍ ẘ w'(±ç) v ʋ
e.g. wind (1), when, while, where (1,2,3,4),
we (=us), will (5), when, which (6), away (7,8)

6. Transcription, analysts and consistency

There is little material on the reliability of transcriptions (whether of a segmental or of a prosodic kind) except, especially, Ladefoged (1960) and Ringaard (1965). It is certainly peculiar - given the phenomenal rise in output of variationist and sociolinguistic research - that so little attention has been paid to variation amongst **analysts** compared to the attention paid to the variation amongst informants. Clearly at least some of the variation attributed to the latter must have its real source in variation arising amongst the former. (One presumes that analysts do not want to be believed to be less, or more, human than their subjects ! (A,C9))

The assumption - for which I argued at length in Chapters 2 and 3 above - that extra-linguistic information is derived from the differences between the acoustic signal and the hearer's experience as a speaker has been embedded in, and interpreted through, a spatial metaphor, and its associations with other (visual) metaphors such as perspective and parallax. If we conceive of any particular hearer as a fixed point in space, then we may base that fixed point upon whatever it is which is used as the basis for comparison, and assume that his perception of the distance of other speakers from himself is a function of their dissimilarity. The basis for the modelling of variation is then a principled search for a general optimum space, or variety space (VSp).

The crucial definition for our purpose now is that of an analytic variety ($_{a}V$) - a profile of linguistic variants which exhaustively places any sample of speech of a particular speaker in the VSp as defined at that moment by its contents and dimensions.

The qualification exhaustively is important since it emphasises that the usefulness of any current version of the VSp is its dimensionality (or criterial properties) together with its contents (or transcribed profiles). Given this definition of usefulness it is necessary to examine the stability of the methods which have been used to select the criteria (dimensionality) and to analyse and code the informants' speech (contents).

What is predicted by the model itself (see Chapters 2 & 3 above) is that different analysts will select different criteria as a direct or indirect result of differences in the analysts' own positions in the space. Furthermore the willingness of analysts to incorporate some criterion must depend, at least in part, on their exposure to it. [N]

[N] I do not feel able to make this claim more definite because of the effects of imagination on the establishing of criteria, especially phonological criteria. One can easily imagine an informant saying something which one has never heard **anyone** saying. For instance in OU {eɪ}, 0106 PDV //ɛ//, I have only **heard** make, take as [mɛk], [tɛk]; nevertheless, [bɛk], bake, may be imagined fairly readily; [rɛk], rake (vb.), requires a good deal more effort, and [lɛk], lake (n.) seems to be somewhere beyond unlikely. Any of these, doubtless, could be produced by a native speaker on the basis of analogy, but under what conditions and with what probabilities I guess that we shall never know (A,C10).

Such limitations on the exhaustiveness of the VSp may be pictured through geometrical properties. Insofar as a given selector under- or over- represents any delimitable subsets of possible criteria, he or she will be operating with a locally deformed version of the VSp. (We may think of a Moebius ring partly sliced along its 'centre' line, or a torus with a bite out of it, or the kinds of compressed and/or expanded graph paper (transformable grids) which Thompson (1942) uses to express time-depth relations amongst skulls, or Sneath (1967) uses to express the relatedness amongst peptide chains.) One might imagine that as the number of investigators increased, and hence as the number of contributory deformations of the dimensionality increased, one would find that their conflation led to a **regular** VSp. That is, that there might be a cancelling effect among such skewings of selection. At present it is impossible even to establish a likelihood for this speculation (A,C11).

It **does** seem to me reasonable to assume, however, that if different analysts have different, but overlapping, preferred sets of criteria with which to handle the speech of a given community, then those different analysts will **transcribe** that speech differently even after periods of communal training. [N]

[N] What I am suggesting is that periods of communal training will not eradicate the different experiences of analysts both as hearers and as speakers, but will overlay them. Between training sessions my guess is that there will always be a drift of their auditory habits away from the communally established norms and back to their 'natural' hearer-speaker experience. This drift will clearly be different for different analysts, it may be different for a single analyst on two separate occasions, and the end 'location', after the drift, may be different as a result of the repeated training (A,C12).

Figure A3 shows the data of transcriptional differences between analysts in respect of spoken text from four informants. Here, three analysts have transcribed the stressed vowel nuclei in the same naturally read passage from four informants. (Some of the following discussion draws on Pellowe et al 1972b.) Three of the informants (notably, A, B, and C) were the analysts themselves.

A comparison of the disparity between each pair of analysts reveals the relative distortions which are likely to affect any hierarchically ordered stressed vowel subspace containing the profiles of surveyed and hand-picked informants.

The disparity index for any pair of analysts **x, y** is given by:

$$\sum_{1}^{m} (p_x - p_y) / m$$

where m is the number of realisations of the particular phoneme being examined, and p the particular diacritic locus given to any cardinal vowel representation of that phoneme by a particular analyst. What counts as a relevant diacritic degree is determined empirically from the data, i.e. the range of a particular diacritic modification for a particular cardinal vowel for a particular analyst.

It is helpful to interpret this material, as I hinted above, by means of perspective differences. In order to do this one needs to know the relative whole-variety 'positions' of analysts. These are given very crudely in terms of a linear scale of localisation in Figure A4. (In part this crudity arises from the fact that the L-varieties of B and C are from different localities.)

Ringaard (1965) sadly concludes that the transcriptions of phoneticians do not tell us so much about the speech of the locations in which they are working as about the speech of the phoneticians themselves. In the face of Fig. A4, and in the context of the model developed above in Chapters 2 & 3, we may focus Ringaard's general point into a premise, notably:

that transcriptional disparities between analysts will be a function of both

(a) the relative differences between the analysts' own positions in the Variety Space, and

(b) the relative differences between the analysts' positions and those of the informants in respect of whose speech their transcriptions are being compared (A,C13). [N]

[N] There is a not quite vicious circle here. It might be argued that we cannot know the positions of the analysts or the informants in the VSp until the analysis has been done, and therefore some such disparities as those we are discussing would already have been included in the VSp contents. Though this is true it does not vitiate the line of argument I am here taking. The idea would, in the end, be to 'calibrate' or 'normalise' the VSp in accordance with the patterns discussed below.

Given this premise our first expectation might be that the disparities for informants B and C should be relatively larger for the analyst pairs A-B and A-C than for the analyst pair B-C, because, according to Fig. A4, the linguistic distances between A & B and between A & C are greater than that between B & C (let us call this 'expectation A').

Figure A5 indicates the reasonableness of this expectation. The expectation is examined for informants B and C separately and for them both taken as a composite (Σ). The x under B opposite /ɩ/

indicates that the disparities between A & B and between A & C are **less** than that between B & C for transcriptions of /ɪ/ in B's speech (expectation contradicted). The <u>✓</u> under Σ opposite /i/ indicates that the disparities between A & B and between A & C are **greater** than that between B & C for transcriptions of /i/ in B's speech **and** in C's speech (expectation confirmed). The <u>?</u> under C opposite /ɛ/ indicates that the disparities between A & B and between A & C are the **same** as that between B & C for transcriptions of /ɛ/ in C's speech (expectation equivocal). The <u>(✓)</u> under C opposite /u/ indicates that **either** the disparity between A & B **or** the disparity between A & C (but not both) is greater than the disparity between B & C in transcriptions of /u/ in C's speech (expectation marginally confirmed).

As can be seen the fate of expectation A is very variable for different stressed vowel phonemes. In numerical terms, overall (column Σ), and applying the weakest definition of 'confirmation' (i.e. ✓ + (✓)), 9 out 14 cases support the general thrust of the expectation. There seems to be no feature common to all those phonemes whose transcriptions strongly - /ɪ/, /a/, /ʊ/, /ɜ/ - or weakly - /eɪ/ - contradict the expectation. It seems as if the transcriptions of informant C provide greater support for the expectation than transcriptions of informant B, which, given the locations in Fig. A4, ought to be the case according to our general premise. However, this conclusion is not very strongly founded if one computes the summed mean disparity per phoneme for all analyst pairs by informant.

For that measure one gets:

X	2.71
A	3.01
C	3.48
B	3.61

Taken as uni-dimensional distances, these figures give roughly the same pattern as Fig. A4, except that the positions of B and C are **reversed**. This finding would seem - contrary to our premise derived from Ringaard above - to indicate that there is something 'difficult' about informant B: since regardless of the analyst pair being compared in B's respect, the disparity for B's stressed vowels overall will be higher than for other informants.

There is an intractibility about this conclusion which one wishes to circumvent, but for whose circumvention there is no obviously promising line of argument.

It might be possible to return to the starting point, to say that there is nothing wrong with the figures just cited but rather to say that what is wrong is the assumption which underlies Fig. A4. (The implicit assumption there was that a crude linear location of the four informants in (impressionistic) whole-variety terms would be sufficiently sensitive to calibrate transcriptional differences of stressed vowel phonemes.)

In other words, given that the following two pictures of informant locations are theoretically and empirically reasonable, we should base our assessment of transcriptional differences in respect of stressed vowel phonemes on the second picture rather than on the first:

(1) locations of informants according to their whole-variety:	X---A------B---C
(2) locations of informants according to their tendency to induce transcriptional disparity:	X---A------C---B

But such a move surely begs more questions than it can answer. The theoretical compatibility of these apparently contradictory locative representations can hardly be attacked either linguistically or computationally.

Linguistically it is not only safer, but more coherent, to assume, following Bazell (1966,1977b - see above Chs. 2 & 3), that the postulate of linguistic solidarity (i.e. between systems) remains to be proved, and that until it is we shall assume non-solidariness amongst those systems. Such an assumption is especially important in the present case, since each 'system' which I am concerned with here is the sum of:

(a) the repertoire of vowel systems of each analyst-as-speaker,

(b) the vowel system of each informant as represented in each informant's spoken output,

(c) the orthographic-realisational representation of the vowels of

(b) by each analyst-as-hearer, and

(d) the orthographic-realisational representation of CVs by each analyst-as-hearer and each analyst-as-speaker.

At best then, each such 'system' is a system of systems; and, at a good deal less than best, each such 'system' is a complex mixture of systems and of individually characteristic, but non-systematic, fluxes.

Computationally we know (see Appendix D below) that many classifications are semi-metric, that is that the representation which shows that A is "close to" B, and B is "close to" C, neither asserts nor denies - is neutral in respect of the possibility - that A is "close to" C.

We shall therefore reject neither of the relationships expressed above, but will behave as if they were in some perfectly natural resonance with each other, albeit a resonance we can neither hear nor make use of for our argumentative ends (A,C14).

In addition to my restrictive expectation (A) that disparities in transcription arise solely from the differences between the analysts themselves, it is reasonable to imagine that some component of the differences is attributable to the difference between the analysts taken as a compound entity, and some particular informant. This more general expectation (B) is that the disparity between analysts increases as the sum of their linguistic distance from the informant increases (A,C15). Figure A6 examines the reasonableness of this expectation in respect of informant X. The expectation predicts that the B-C disparity will be greater than the A-C one, that the A-C disparity will be greater than the A-B one (and hence that the B-C disparity will be greater than the A-B one). If we interpret the prediction as being in respect of all the phonemes analysed then the expectation is more unreasonable than reasonable.

Where should one turn in order to try to save the conjecture (B) from our discontinued entertainment of it ? Perhaps there is some additional information which could be used to constrain or modify the disparity indices. It is possible that part of the disparity arises neither from differences between the habits of the analysts as speaker-hearers in their own lives, nor from the differences between them and the informant whose speech they are analysing, but from different reference norms or transcriptional conventions. Some of the difference may arise from the technology (rather than the technique) of the transcriptional process.

One analyst's [i̠] might not be the same as another's either because CV 1, [i], may be articulated differently by the two (deviant reference norm), or because of a wider or narrower range of diacritic usage, or both.

In order to be able to consider this possibility, each of the analysts (A, B, C) recorded a randomised series of Cardinal Vowels which were analysed by all of the analysts including the performer. Disparity indices for pairs of analysts as both performers and hearers are given in Figure A7. Thus, the first column gives the disparities between analyst A and analyst B in respect of CV realisations performed by A, whereas the second column gives the disparities between analyst A and analyst B in respect of CV realisations by B. The pair mean (rounded down to whole numbers) is given in the third column.

Several interesting facts emerge from this table:

(1)

The greatest mean analyst-pair disparity, across all cardinal vowels, is between B and C (2.55), the second greatest is between A and B (2.20), and the least is between A and C (2.16). This is true whether one considers all CVs, as above, or just CVs 1-8 (in which case the figures are 2.55, 1.50, and 1.37 respectively).

In the light of our expectations expressed in Figs. A5 and A6 above, this seems somewhat surprising, since B and C have the most close whole-variety positions amongst the three analysts, and A and C have the most distant (Fig. A4). [N]

[N] One reason why those expectations **may** not be applicable in this situation is that CVs are patently not part of any variety of English. And therefore the usual dichotomies and distances, which separate hearers who themselves have different varieties, are in abeyance. Dichotomies such as in-group/out-group, often-heard/little heard, stereotyped/non-stereotyped cannot apply - presumably - to non-English vowels heard in isolation.

---

(2)

Beneath this regularity in mean analyst-pair disparity (Fig. A7), however, there are important differences between the pairs of analysts, and between the performers.

(a) Analysts A and B seem to have very similar disparities whether the vowels being analysed are those produced by A (2.20) or those produced by B (2.21). (But the breakdown of means on all CVs against CVs 1-8 only shows that this similarity is not evenly distributed across the vowels, since there is a mean A-B disparity of 1.25 for A's CVs 1-8, against 1.75 for B's.)

(b) By apparent contrast, the disparity between analyst A and analyst C when they are analysing C's CVs 1-16 (2.73) is more than two thirds bigger than the disparity when they are analysing A's (1.60). A difference of a similar kind, though not so striking, exists between the disparities between B and C when they are analysing C's CVs 1-16 (2.80), and when they are analysing B's (2.28).

What these facts seem to indicate is that the disparities between analysts are not singly, or simply, or even primarily, a reflection of the differences between them as speakers of their own varieties. But these disparities are, rather, reflections both of some characteristic of the performed CVs themselves, and of different auditory/perceptual characteristics of the CVs which are picked out or even preferred by the different analysts. We may focus on this by examining the individual pair means for CVs 1-8 only (Fig. A7). It is notable that the lowest disparities are between analysts A-B (1.25) and analysts A-C (1.12) when they are analysing A's CVs.

In other words A's CVs seem to elicit auditory agreement amongst pairs of analysts. Not only this, but, if the **range** of disparity is taken into account (1.25 - 1.12 = 0.13), and if we take a low range of disparity to indicate agreement to agree, or agreement to disagree) then A's CVs seem to elicit agreement to agree.

B's CVs are very different in this respect, if we consider the B-A and B-C disparities. Of the six pairwise disparities, B's CVs are responsible for the two highest (B-A = 1.75; B-C = 2.37); and, in addition, the range between these disparities (0.62) is greater than it is for either of the other performers (A, C). That is, B's CVs elicit auditory disagreement, **and** they elicit disagreement about that disagreement.

C's CVs are intermediate between those of A and those of B in both these respects - the level of auditory agreement is intermediate (C-A = 1.62; C-B = 2.12), and the degree of agreement about this level of agreement is intermediate (range = 0.50).

So much for the performers. What about the analysts ? The most notable single fact is the degree of difference between all the analyst-pair disparities (regardless of whose CVs are being analysed) in which A is one of the analysts, and those analyst-pair disparities in which A is not one of the analysts. It is difficult to imagine what process could be at work here. (It certainly looks accommodative, though even I am at a loss for a speculation, however implausible.)

But we may be able to construct a parallel line of argument as follows. Even though the disparities between B-C and C-B are the highest of the three pair-wise sets, the range of those disparities is the lowest of the pairwise sets (0.25). We might, apart from the collusive implication, say that this constituted an agreement to disagree about each other's CVs. This now begins to look like something which may have nothing to do with CV realisations as such, but with something as distant as paralinguistic characteristics, or, even, personality. What I am suggesting is that B's perception of C's CVs and C's perception of B's CVs may share in common a skewing which arises from their varietal paralinguistic norms as ordinary speakers (or, if you will, the inferred equivalents in terms of the whole person). This may be rendered more plausible if one admits the possibility that, for speakers of different localised varieties such as B and C, CVs in general may be associated with the paralinguistic norms of non-localised varieties (RP). (This would certainly fit with the facts about A's CVs, cited above.)

Let us return now to a consideration of the data of Fig. A3 (the original data) in condensed form. Figure A8 presents the **summed** mean disparity indices, by pairs of analysts, for the stressed vowel phonemes of the four informants. In what way can we constrain these measures using the facts of Fig. A7 (the CV data) ? We could, simply, subtract the mean pairwise CV disparities from the real data disparities of Fig. A8 - this is done in Figure A9. But clearly all that such a process will do is to reduce the general levels of disparity without affecting any rankings (amongst either the analyst-pairs or the informants). It will not have the effect which is wanted, which is to equalise, more or less, the disparities both amongst analyst-pairs and amongst informants.

How then should one proceed ? As I implied above, comparing the range between the mean disparities for each pair of analysts on their two sets of CVs (see below) seems almost to encompass both the facts we have noted about the performance of the CVs and the whole-variety positions of the analysts. This is certainly true when we consider the disparity range across all CVs; but the pattern is typologically different for CVs 1-8 only. Perhaps incorporating such extra facts in a calibrational index might improve on the simple subtractive principle of Fig. A9. ?

	Range of CV disparity	
	All CVs	CVs 1-8
A-B	0.01	0.50
B-C	0.52	0.25
A-C	1.13	0.50

It looks as if consideration of range alone is not going to provide us with a straightforward method of calibrating analysts. Perhaps if we compare the overall mean disparities with the ranges of disparity for each analyst pair we might find that something useable emerged? We certainly see that the rankings on each measure are different:

	All CVs		CVs 1-8	
	range	mean	range	mean
A-B	0.01	2.20	0.50	1.50
B-C	0.52	2.55	0.25	2.25
A-C	1.13	2.16	0.50	1.37

If either of these measures - the range of disparity, or the overall mean disparity - is an optimum one for constraining the variability amongst analysts it is not at all clear which of them it is (A,C16). Given this uncertainty, it might be reasonable to construct some additive or multiplicative combined index from them and to hope that this would constitute a useable compromise. In the two cases we get:

	additive index of range and mean		multiplicative index of range and mean	
	All CVs	CVs 1-8	All CVs	CVs 1-8
A-B	2.21	2.00	0.02	0.75
B-C	3.07	2.50	1.37	0.56
A-C	3.29	1.87	2.44	0.68

It is clear that there can be, in the present state of the art, no principled basis for deciding between an additive index and a multiplicative index for constraining the disparities between analysts. Indeed it is not entirely obvious why one should choose one or other of these two, in preference to, say, a simple index based on either the mean or the range of disparities. The only argument in favour of a complex index is that both the range and the mean disparity appear to tell us something useful about pairs of analysts.

If the index we choose is the additive one (for CVs 1-8) then it will have to be used divisively (since those indices are greater than unity); if the index we choose is the multiplicative one (for CVs 1-8) then it will have to be used multiplicatively (since the indices are less than unity).

With the present amount of material, however, it seems to me that one can get very little further in **deciding** upon an appropriate strategy. It is enough, as far as I am concerned, to have seen what the problems are, and what sorts of solutions are going to have to be found.

The contents of any VSp, whether that VSp is computerised, imagined, or the inferential basis for some particular hearer's behaviour, are products not only of the dimensions of that VSp and of the sampled varieties which are dispersed upon those dimensions. But such contents are also, if not equally, the products of the perceptual processes of the human analysts who are involved in creating that VSp. In so far as more than one analyst is involved, just so far will it be the case that differences between the analysts will be silently incorporated into the VSp. If these differential contributions remain silent, then the VSp will be distorted in ways which will not even be seen. Clearly the problem has not been solved here, but it has been noticed.

(Cf. Kerswill & Wright (1989), especially their last paragraph

> the problem lies in an inherent multi-layered ambiguity in the task of transcription itself. [It] represents articulations, or . . . auditory impressions; . . . discrete segments . . . [presupposing] a prior phonological analysis, or . . . a continuously varying acoustic signal. . . . The snag is, all these things are true to different degrees, and unfortunately the transcribers will put the boundary between each of the pairs of opposites in different places. . . . Transcibing without any kind of theory is a dangerous thing: we simply do not know exactly what each individual is doing, and consequently we cannot interpret precisely what they write down (1989: 57-8).)

7. Classificatory fate of criteria.

Here, very briefly, I shall consider the fate of the criteria specified in Sections 3, 4, & 5 above when classifications of speakers are constructed. (Much more detail can be got about the same matters from Jones (1978 (= Jones-Sargent 1983)) and Pellowe & Jones (1978b, 1979b).)

It is important to stress that, given the profiles of some sample of speakers analysed in respect of some set of criteria, many different possible classifications can be constructed upon those profiles. And deciding which of those classifications is best is not always easy, or even possible. (For further discussion of these matters see above, Chapter 4, and below, Appx. D.)

Here however we shall concentrate upon how the criteria (dimensions) behave in respect of each other under the imposition of a sample of speakers.

The sample is of 52 speakers. These speakers are a mixture of speakers from two random samples, one drawn in Gateshead-on-Tyne (the continuation of the conurbation south of the river), and one drawn in Newcastle-on-Tyne (north of the river). The methods by which the random samples were drawn was different in the two cases, and there is no statistical argument here being advanced as to why such added samples <u>should</u> be statistically respectable. (We shall simply eschew making any inferences about population values from sample statistics.) The point of the present review is simply to examine what happens to the segmental criteria we have just specified.

The phonological realisations of each informant are represented by a stream of numerical codes, each of which specifies the state of the PDV of the OU which some particular sound represents in some particular word. A series of computer programs written by Val Jones (for full specifications see Jones (1978)) counts the raw frequencies of each state for each informant, converts these raw state frequencies to within-OU percentages, calculates sample statistics for variables, and deletes variables with zero variance since these cause difficulty for many classificatory methods (Everitt 1974). These data were input to CLUSTAN 1A, which is a suite of classificatory programs designed by Wishart (1969), in the form of a matrix. (A matrix of informants-by-criteria for a classification of informants; a matrix of criteria-by-informants for a classification of criteria.) This classification is based on 122,000 segmental realisations in the speech of the 52 informants. After the elimination of zero-variance criteria these realisations are distributed across 542 states representing 51 OUs. But the maximum number of criteria which can be handled in a single CLUSTAN 1A run is 200, therefore the phonological criteria were divided into three phonological subspaces as follows:

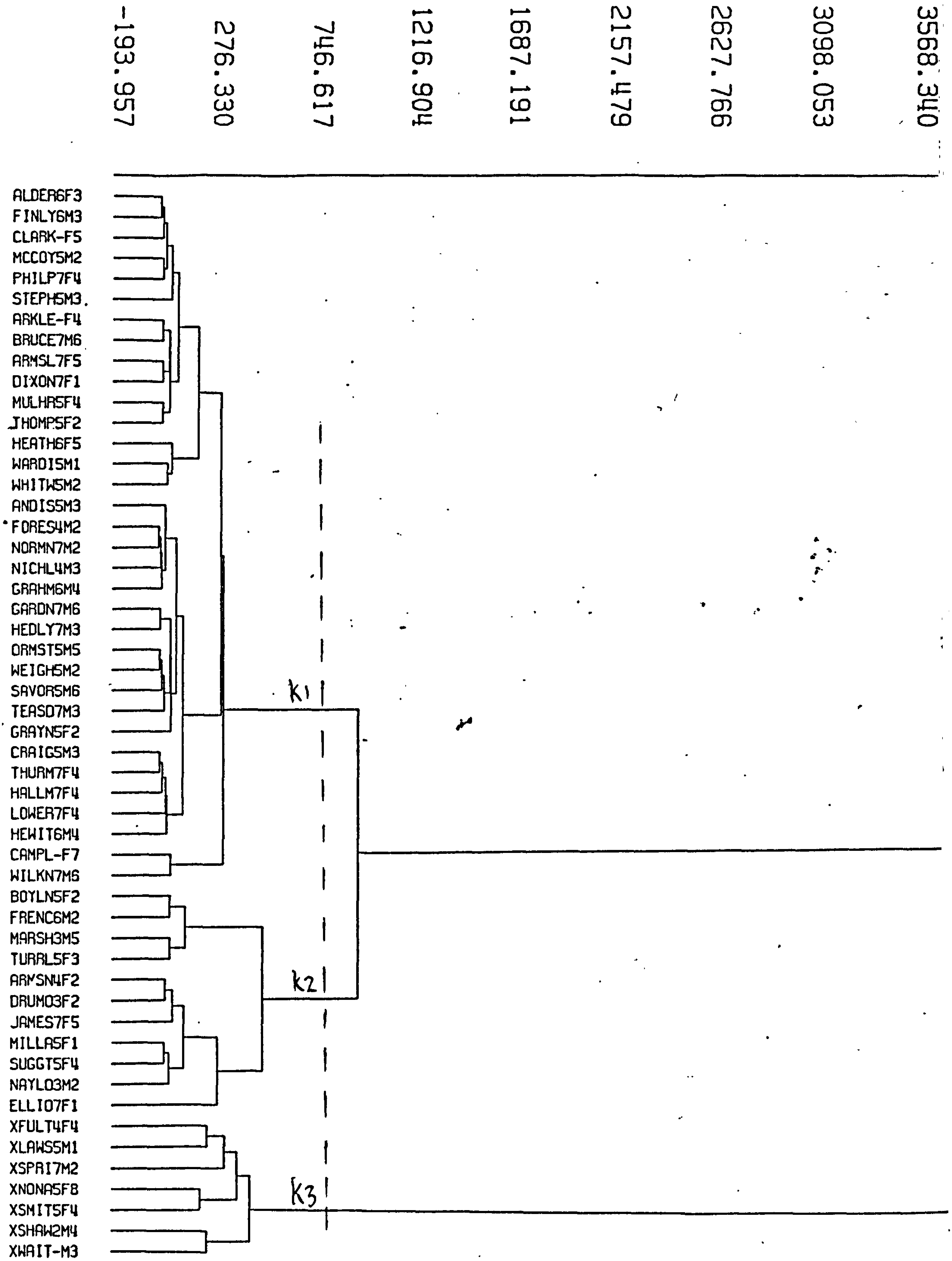
Name	**OUs**	**Numbers**
%FON1	{iː}{ɪ}{ɛ}{a}{ɑ}{ɒ}{ɔ}{ʌ}{ʊ}{u}	10 OUs, [154 Sts]
%FON2	{eɪ}{əʊ}{aɪ}{aɪə}{aʊ}{ɔɪ}{ɜ}{ɪə}{ɛə}{ʊə} {ə$_{3}$}{ə$_{4a}$}{ə$_{2,1}$}{ɪ$_{1}$}{ɪ$_{2}$}{ə$_{4b}$}	16 OUs, [189 Sts]
%FON3	{p}{b}{t}{d}{k}{g}{tʃ}{dʒ}{f}{v}{θ}{ð}{s}{z}{ʃ}{ʒ} {h}{m}{n}{ŋ}{l}{r}{j}{w}{*ɪŋ}	25 OUs, [199 Sts]

Thus, %FON1 comprises the lexical sets for monophthongs (though the realisations need not be monophthongal), %FON2 comprises the lexical sets for diphthongs and reduced vowels (though the realisations may not be diphthongal or reduced respectively), and %FON3 comprises the lexical sets for consonants. The next three pages show dendrograms for the classifications of the 52 informants in the three subspaces of %FON1, %FON2, and %FON3. The classifications took the following options in CLUSTAN 1A: the comparison function which was used was Euclidian D^2 (squared distance), and the clustering method was Ward's, which is a hierarchic agglomerative method based on minimising the sum of squares. (Detailed discussion of comparison functions and clustering methods, amongst other things taximetric, may be found in Appx. D, below.) Dissimilarity (in the form of squared distance) can be read off the scale on the left, the five-letter codes are mnemonics for informants. Perusal of the dendrograms tells us:

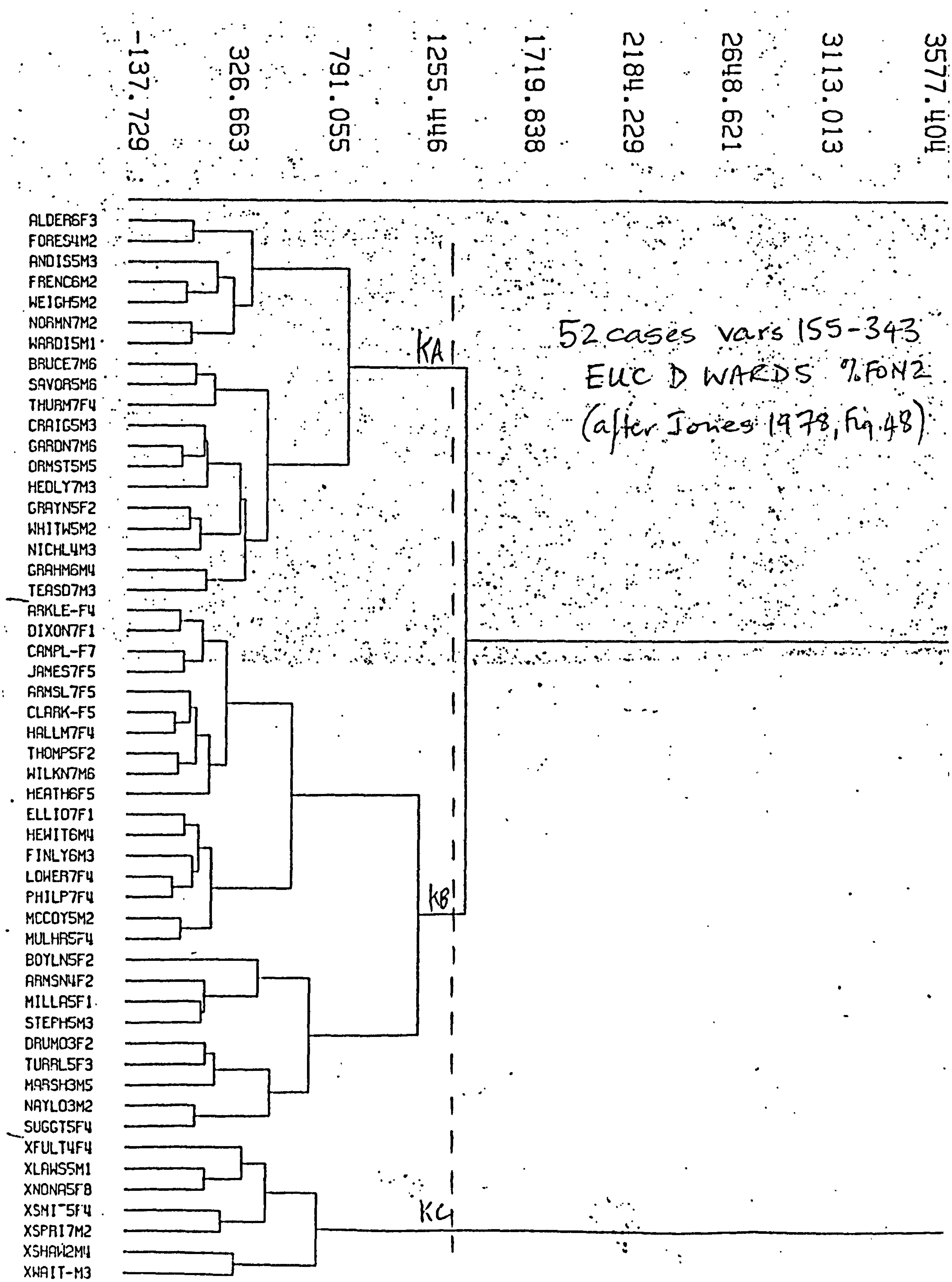
(a) that informants are clustered differently in the three subspaces in two respects:

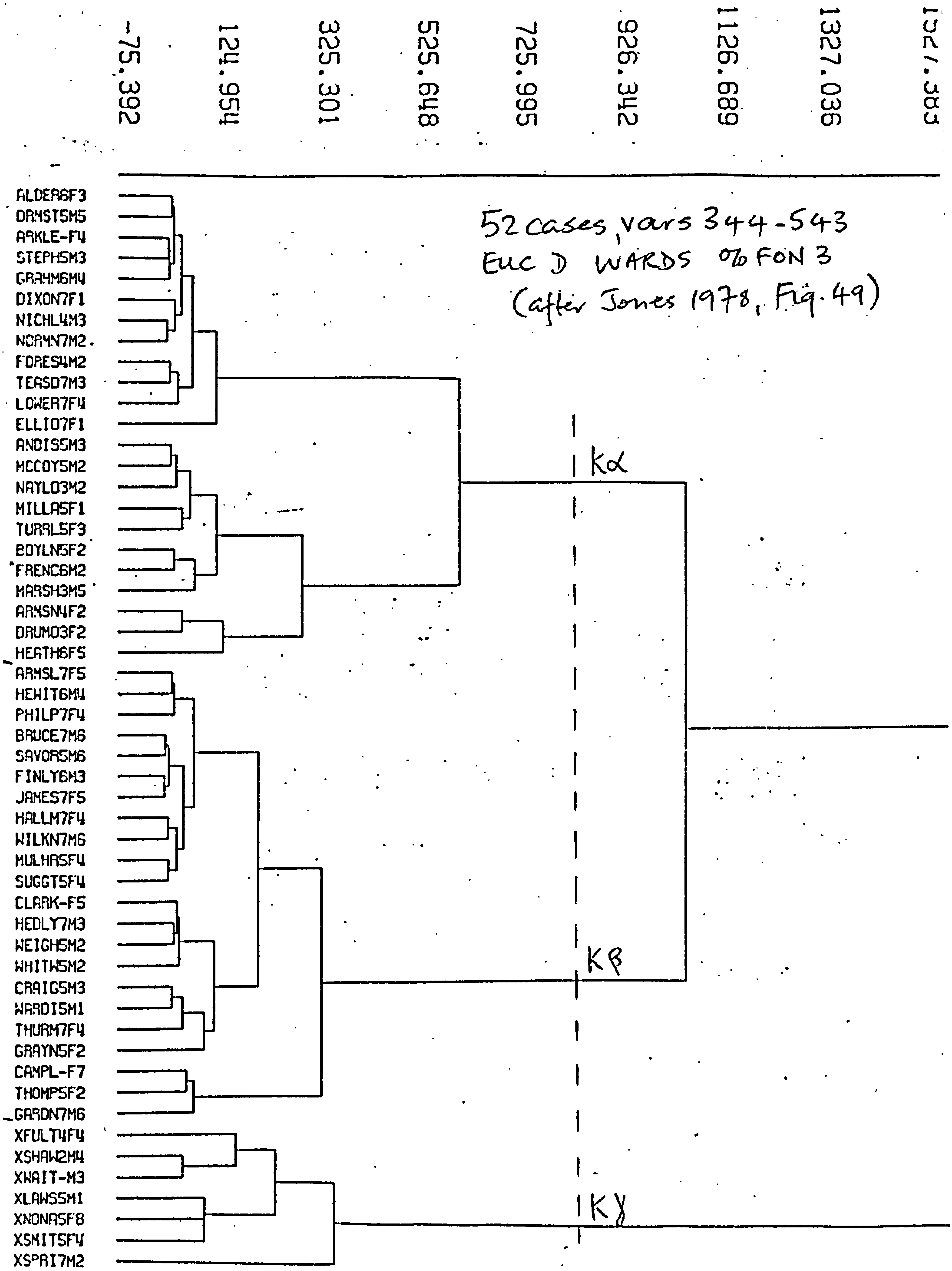
(i) the order of their joining the groups is different, and

(ii) the overall membership of the groups alters as between the three subspaces;



52 CASES VARS 1-154 %FON1 EUC D WARDS (after Jones 1978, Fig. 4.7)


-137.729
326.663
791.055
1255.446
1719.838
2184.229
2648.621
3113.013
3577.404
52 cases vars 155-343
EUC D WARDS %FON2
(after Jones 1978, Fig.48)
KA
KB
KC
ALDER6F3
FORES4M2
ANDIS5M3
FRENC6M2
WEIGH5M2
NORMN7M2
WARDI5M1
BRUCE7M6
SAVOR5M6
THURM7F4
CRAIG5M3
GARDN7M6
ORMST5M5
HEDLY7M3
GRAYN5F2
WHITW5M2
NICHL4M3
GRAHM6M4
TEASD7M3
ARKLE-F4
DIXON7F1
CAMPL-F7
JAMES7F5
ARMSL7F5
CLARK-F5
HALLM7F4
THOMP5F2
WILKN7M6
HEATH6F5
ELLIO7F1
HEWIT6M4
FINLY6M3
LOWER7F4
PHILP7F4
MCCOY5M2
MULHA5F4
BOYLN5F2
ARMSN4F2
MILLA5F1
STEPH5M3
DRUMO3F2
TURAL5F3
MARSH3M5
NAYLO3M2
SUGGT5F4
XFULT4F4
XLAWS5M1
XNONA5F8
XSMI¯5F4
XSPRI7M2
XSHAW2M4
XWAIT-M3


-75.392
124.954
325.301
525.648
725.995
926.342
1126.689
1327.036
52 cases, vars 344-543
Euc D WARDS % FON 3
(after Jones 1978, Fig. 49)
Kα
Kβ
Kγ
ALDER6F3
ORMST5M5
ARKLE-F4
STEPH5M3
GRAHM6M4
DIXON7F1
NICHL4M3
NORMN7M2
FORES4M2
TERSO7M3
LOWER7F4
ELLIO7F1
ANDIS5M3
MCCOY5M2
NAYLO3M2
MILLA5F1
TURAL5F3
BOYLN5F2
FRENC6M2
MARSH3M5
ARMSN4F2
DRUMO3F2
HEATH6F5
ARMSL7F5
HEWIT6M4
PHILP7F4
BRUCE7M6
SAVOR5M6
FINLY6M3
JAMES7F5
HALLM7F4
WILKN7M6
MULHA5F4
SUGGT5F4
CLARK-F5
HEDLY7M3
WEIGH5M2
WHITW5M2
CRAIG5M3
WARDI5M1
THURM7F4
GRAYN5F2
CAMPL-F7
THOMP5F2
GARDN7M6
XFULT4F4
XSHAW2M4
XWAIT-M3
XLAWS5M1
XNONA5F8
XSMIT5F4
XSPRI7M2

(b) the different subspaces have differential divisive power on the sample, as follows:

	First fusion	Last fusion
%FON1	19	4243
%FON2	73	4255
%FON3	17	1855

The fact that speakers in the same speech community (having roughly the 'same' accent) should be differentially similar to one another in respect of variables from subspaces of the 'same' linguistic system should not be surprising (see above, Chapter 4, for discussion of Bazell's postulate of non-solidarity). But it is a type of fact not of widespread currency in sociolinguistics. As Jones (1983: 195-6) has it

"If the three classifications produce similar distributions of informants across all clusters, then it will be demonstrated that [some] subset of variables [would be] an adequate basis for representing linguistic variability, and an exhaustive inclusion of variables [would] mean the inclusion of redundant information. If, however, the sample clusters differently with respect to the three subsets of variables, we can say, with confidence, that the subsets of variables chosen only produce partial classifications, and cannot be taken as representing overall . . . segmental phonological variability."

Note that although the consonantal subspace (%FON3) is less discriminating than the vocalic subspaces, it does discriminate groups (hence must contain significant variation) and constitutes them differently from the other two subspaces.

It will be noted that in all three subspaces, we have chosen to consider the sample as having been divided into three clusters (the 3K level, K standing for 'cluster'). But, it may be asked, given that there is a continuous reduction in the numbers of clusters for increasing values of D^2, why should three clusters have been chosen as best representing the sample in the three subspaces ? If the values of D^2 are plotted against the number of clusters for each of the subspaces (%FON1, %FON2, %FON3), it is found that the largest single plateau of unchanging D^2 values occurs at K = 3 in %FON1 and %FON3, but at K = 4 in %FON2. Nevertheless, in order to facilitate comparison, we chose to represent the sample at the 3K level in all three subspaces. (In addition, though of no relevance to our present purpose, the best representation of the same sample in the social classification was also at the 3K level.)

Figures A10, A11, A12, and A13 show how the sample members are differentially distributed with respect to three subspaces. (Figures A10 - A17 are in a pocket at the back of Vol. iii.) Figure A10 shows the clumped sample members by mnemonic. Since the clustering is shown topologically for each subspace, each sample member remains in the same place for each figure.

Overlaying Fig. A10 by Figure A11 will give the distribution of the sample in %FON1 (the 'monophthongal' subspace), whose Ks are labelled with numerals (K1, K2, K3). Overlaying Fig. A10 by Figure A12 will give the distribution of the sample in %FON2 (the 'diphthongal' subspace), whose Ks are labelled with roman capitals (KA, KB, KC). Overlaying Fig. A10 by Figure A13 will give the distribution of the sample in %FON3 (the 'consonantal' subspace), whose Ks are labelled with Greek letters $K\alpha$, $K\beta$, $K\gamma$.

In the three subspaces, one cluster maintains its membership, though not, as specified above, its levels of relationship - K3 = KC = $K\gamma$. Four possible explanations of this spring to mind:

(a) these seven, X-prefixed, informants are the only ones from Newcastle, the rest being from Gateshead;

(b) the sampling method for the Newcastle informants was different from that used for the Gateshead informants;

(c) one analyst analysed all the X-prefixed informants, and a different analyst analysed all the other informants;

(d) a mixture of all three or any two of these differences.

It is impossible on the available evidence to decide which of these four is the likeliest source of this cluster maintenance; but it seems reasonable to me to conjecture that there is something abnormal about this pattern.

For the rest of the clusters - K1, K2, KA, KB, Kα, and Kβ - the distributional relationships across the three subspaces (in terms of numbers of informants) are as follows:

	KA	KB
K1 (34)	18	16
K2 (11)	3	8

	Kα	Kβ
K1 (34)	14	20
K2 (11)	9	2

	Kα	Kβ
KA (21)	9	12
KB (24)	14	10

What is entirely clear from these representations is that the three subspaces represent the sample in radically different fashions. There can, surely, in the light of such complex dependencies be absolutely no justification for excluding any, let alone a majority, of these criteria. Such exclusion certainly cannot be undertaken on any reasoned grounds, for why should the reasoning start in one place rather than another ?

We can move further along the same line of conclusion by considering Figure A14, which shows the 'derived' clusters which arise from the superimposition of Figs. A10, A11, A12, A13. Here we find that the sample is divided into eight groups. Members of any given group share the same grouping in all three subspaces and they are thus, if only in a secondary sense, monothetic or 'Aristotelian' groups (i.e having necessary and sufficient characteristics for group membership).

<u>None</u> of these eight groupings could be deduced from any pair of subspace groupings (and this <u>includes</u> any pair of K3/KC/Kγ). This seems to provide disastrous counterevidence against, amongst other

things, anything like an implicational scale.

The fate of the criteria in these three clusters are as follows. The figures represent the cluster mean frequencies of the most frequent PDVs of each OU, by cluster. Frequencies of less than 10% are not shown.

		Mean frequency in %FON1		
OU	**PDVs**	**K1**	**K2**	**K3**
iː	iː	60	69	54
	ɯi	31	20	33
ɪ	ɪ	60	67	83
	ɐ	25	14	
	ɜː		14	
ɛ	ɛ	82	83	81
a	a	86	81	84
ɑ	ɑ	92	77	84
ɒ	ɒ	92	95	87
ɔ	ɔ	70	69	84
ʌ	ɷ	62	35	26
	ɪ	18		
	ə		39	45

OU	PDVs	K1	K2	K3
ω	ω	88	64	46
	u			27
u	u	89	85	90

It is notable that the discrimination between K1, K2, and K3 (i.e. amongst monophthongal criteria) hardly rests at all on differences at the level of PDV, but rather depends upon qualitative <u>and</u> quantitative differences at the level of Sts. These differences will not be explored further here (for further details see Jones (1978, 1983), Tables 13-25). And, where there <u>are</u> differences which show themselves at PDV level (in OUs {iː}{ɪ}{ʌ} and {ω}), it is clear that in the last three cases the analyst of the informants in K3, and the analyst of the informants in K1 and K2 never **both** contribute to the appearance of these differences. [N]

[N] The analyst of the informants in K3 is the same as analyst B in the calibration experiment of Section 6, above. The analyst of the informants in K1 and K2 is the same as analyst C in that experiment.

The criteria of %FON2 show rather different patterns. There is much more variability at PDV level than there is in %FON1. The cluster mean frequencies of PDVs in the %FON2 clusters are:

		K mean frequency in %FON2		
OU	**PDVs**	**KA**	**KB**	**KC**
eɩ	eɩ		5	28
	iː	61	78	5
	iə	30	4	26
	ɛɩ			16
əɷ	əɷ	1	6	67
	ɔː	29	73	4
	uː	24	5	20
	ɑː	38	4	
aɩ	aɩ	5	12	42
	ɑː	52	56	43
	eɩ	40	25	3
aɷ	aɷ	4	22	32
	ɛɷ	25	60	58
	ɔɷ	51	6	
ɔɩ	ɔɩ	1	20	7
	ɷɩ	50	48	21
	əɩ	25	9	
ɜ	ɜː	1	7	65
	ø	39	63	13
	ɛ(ə)	20	12	12
	ɔ	15	5	
ɩə	ɩə	1	9	29
	iː	24	24	3
	iɛ	71	57	59

OU	PDV	KA	KB	KC
εə	εə	1	1	55
	ε	97	97	29
	ɜː			12
ə$_3$	reduced	21	26	97
	nonred.	76	71	1
ə$_{4a}$	red.	12	40	99
	nonred.	65	43	1
ə$_2$ə$_1$	red.	78	79	18
	nonred.	20	18	
ɩ$_1$	red.	93	90	29
	nonred.	5	8	69
ɩ$_2$	N/A			

Criteria which are diagnostic will be specified as such only insofar as they are diagnostic in <u>two</u> senses. Firstly they must be diagnostic in terms of the definition of that term used by the CLUSTAN 1A suite of programs; that is, an F-ratio of < 1, and a T-value of $> Ø$.[N]

[N] CLUSTAN 1A (Wishart 1969) provides F-ratios, T-values, cluster means and standard deviations for all non-zero-variance variables in every run. The F-ratio is defined as

$F_j = S_{cj} / S_j$, where S_j is the standard deviation of the

j^{th} variable, and S_{cj} is the S.D. of the j^{th} variable in cluster **c**.
The T-value is defined as

$T_j = (\overline{X}_{cj} - \overline{X}_j) / S_j$, where $\overline{X}_{cj}$ is the cluster mean value for variable j in cluster c.
F-ratios < 1 indicate that the within-K variance is lower than the sample variance for the variable in question. T-values which are ≠ Ø indicate that the within-K mean for that variable differs from the sample mean value - positive deviations indicate a higher within-K mean frequency, negative deviations a lower within-K frequency. (Jones (1978, 1983) gives tables of F-ratios and T-values for the criteria, but as I have suggested above, they are difficult if not impossible to interpret reliably.)

The problem about taking these measures seriously is that they assume complete independence between the 'variables' which, in the present case it will be recalled, are Sts. Now even though <u>we</u> know that these Sts are being calculated on a within-OU % basis (and that this form of treatment will still represent a ghosting of the PDV structure which intervenes between Sts and OUs), the CLUSTAN 1A programs behave as if every single St. of the 154 Sts in %FON1 are independent of each other.

Therefore, secondly, we require, for something to be considered diagnostic (whether it be a St or PDV score), that it discriminate <u>all three clusters</u> (KA, KB, KC) from each other. Under both these definitions of diagnostic, the %FON2 subspace yields six OUs which conjointly satisfy them as follows:

{eɩ}{aɩ}{əɷ}{ɜ}{$ə_{4a}$}{$ə_{3}$}.

I shall not, here, present the %FON3 statistics, since most of the criteria in that subspace are effectively at the St level. (The relevant tables and commentary may be found in Jones (1978,(1983), Tables 45-57.) Other matters arising from this classification have been dealt with in Chapter 4 above. For the social classification of this sample see below, Appx. E, p. 558A ff.

Appendix B

Establishing prosodic criteria

Transcription conventions

* onset
/ pitch prominent syllable
F fall
R rise
L level
FR fall-rise
RF rise-fall
F+...........+R (FpR in text)
R+...........+F (RpF in text)
[these representations of tones all immediately precede their nuclear syllables in transcriptions of text]
ehb extra high booster
hb high booster
b booster
c continuance
d drop
ld low drop
absence of any code implies Ø in respect of p-r or tone

A. Getting criteria: problems of principle.

We may assume that it has been known for a very long time by hearers, both trained and untrained, that intonational usage varies socially and geographically (B,C1). Thus Skeat (1911:1-2) writes

> "such differences [between standard and local dialect speakers in respect of pronunciation] are especially noticeable in the use of vowels and diphthongs, and in the mode of intonation."

Regrettably, Skeat gives us no details. Hughes, however, in his vital use of primary eighteenth century sources, not only gives us an idea of the longevity of North Eastern intonation patterns, but also a picture of the ways in which the parents of the children who used those patterns tried to displace them:

> "In 1748 . . . Mr. Carr of Whitworth sent his younger son, Ralphy, 'to a very private school at Craike to wear off the Newcastle tone' " (Hughes 1952: 363).

Another characterisation also gives rather precise information about the distribution and phonetic quality of the 'Newcastle tone':

> "[In 1790, Mr. Ellison] animadverted on [his] boys' accent [having sent them to school near High Wycombe from Gateshead] - apparently the north country tendency to raise the voice on the

last syllable had not been eradicated. The headmaster replied: 'It was impossible their manner of Reading should have escaped you; it has long been a subject of conversation here, sometimes of mirth, at other times we have treated it very seriously, particularly to Master Ellison. He can inform you that he has hardly ever said a lesson or read an English book to me without my talking a great deal to him about it. I have only observed that he generally spoke the last syllable in a sentence nearly a third above the last but one. I have made him repeat the concluding syllables after me and have sunk my voice which he exactly imitated and therefore doubt not but we shall acquire a proper cadence in time' " (Hughes 1952: 365). [N]

[N] I am grateful to Barbara Strang for drawing my attention to Skeat's remark, and to the material in Hughes.

In spite of such kinds of anecdotal knowledge of regional and social variation in the realisation of intonational systems, methods do not exist for making such knowledge precise. The primary problem in establishing such methods is establishing what the variables are. I now examine the reasons for this being the case.

First, however, we may dispose of a difficulty which in my opinion is more apparent than real. Labov (1966) presents four criteria whose conjoint satisfaction he believes will determine useable variables for quantitative work in sociolinguistics. Such putative variables must be:

(i) frequent in occurrence,

(ii) integrated with some higher level of linguistic structure,

(iii) not amenable to conscious suppression by speakers,

(iv) salient.[N]

[N] There are both logical and operational difficulties with these criteria. (i) The required order of frequency is nowhere specified, and it is difficult to imagine by what methods it could be. Furthermore, it is unclear whether the criterion - of frequent occurrence - is introduced for the purpose of methodological economy (e.g. 'too much' data, or distortive methods, are required for handling 'infrequent' variables); [for arguments against this possibility see below, where it's argued that the combining of simple criteria into complex dimensions will reveal more pattern (Bateson 1979b)]; or for unspecified theoretical reasons (e.g. inferences based on 'frequent' variables are intrinsically more trustworthy than those based on 'infrequent' variables; [for arguments against this possibility see above, Ch. 3, where the rarity problem & Preston's veil-line are discussed]. (ii) The meaning of 'integration' is opaque. The criterion may or may not be hiding an important conjecture about solidarity between linguistic systems, but as I have argued above (Ch.s 2 & 3 & cf. Bazell 1949:116) <u>just</u> such conjectures should be at the heart of sociolinguistic effort. (iii) What 'conscious suppression' is, and how it is to be recognised or measured is nowhere described. Are there <u>degrees</u> of amenability to the phenomenon in different types of linguistic item ? How are these to be specified ? On the surface of it criteria (iii) and (iv) appear to be in conflict (even if the former is a 'speaker criterion' and the latter a 'listener criterion', which we don't know anyway). Certainly both hide conjectures behind their axiomatic form (Lakatos 1976: 142), begging such important sociolinguistic questions as: What conditions of interacting promote (or depress) a speaker's awareness of <u>which</u> realisations from <u>which</u> systems whilst she is speaking? Do these conditions <u>always</u> produce this awareness ? Do speakers behaving in this manner have a construction of what their behaviour signifies ? - such that, for instance, to perceive such behaviour in another would lead them to attribute the 'same meaning' to it as to their own ? Is a feature which is 'salient' for (?of) a speaker in one interaction, salient in all other interactions for (?of) that speaker ?

Labov believes, rightly I am sure, that intonational features certainly satisfy his criteria, but he rejects them as suitable variables. He does this on the grounds that " at present we lack the large body of theory and practice in codifying intonation which we have for segmental phones " (1964: 176).

I find this rejection peculiar on two grounds, one pragmatic and one logical. First, how large a body of theory and practice in codifying intonation is necessary before intonational variables can be studied in terms of their variation ? [N]

[N] I assume that 'codifying' means 'coding', or in more common linguistic parlance 'transcribing'.

If the purpose of establishing variables is to divide up (some sample of) the speech community in respect of some non-linguistic characteristics of that community, then it suffices that the variables can be analysed in the speech collected from that community. To be sure, there are several different methods available, and there is also some understanding of the ways in which analysis is neither easy nor, strictly, determinate (Crystal

& Quirk 1964: 32 fn. 1 & passim; and see below, Appx. B, section F), but then, both points are similarly applicable to segmental phonology (Ladefoged 1960; Lyons 1962; Bazell 1966; and see above, Appx. A, 6) for which there __is__ 'a large body of theory and practice'. All that we need to be able to be sure of is that the output of each speaker can be represented equally efficiently by whichever method of analysis we choose to adopt. [N]

[N] In my study of method so far, "equally efficient representation" of speakers (see above, passim) is a straightforward requirement. However it might be expected to run into difficulties as far as the most egregious members of some sample were concerned. Given that some method of analysis has been developed largely in terms of a particular variety, __n__, it might be the case that its representational power in respect of some other variety, __m__, was weaker. (There are a number of ways in which this might be true, but one example will suffice. Imagine that all the categories of analysis which were applicable to __n__ might, under the same definitions, be applicable to __m__ with the exception of two, say $\underline{c}_1$, $\underline{c}_2$, the distinction between which must be redefined to capture relevant facts of __m__. In this case the capacity of the method to represent __m__ is weaker than its capacity to represent __n__. Problems analogous to this are dealt with above (Appx. A,1) and below (Appx. B, section E).

Secondly, the logical peculiarity about Labov's rejection of intonational variables derives from his own firm conviction (Labov 1966) that the motivation for sociolinguistic work is the solution of problems central to linguistic theory. (Rather, that is, than the 'amassing of data from more and more locations using a standard set of methods'.) Two paradoxes are implicit here. In the first place, problems of theory can barely be imagined unless the 'bodies of practice', which are one of their sources, are lumpy, or patchy, or inadequate in some way (cf. Postal (1968), in this respect, on the relationship between systematic & autonomous phonemes). In the second place, in refusing to develop relevant practice and theory in, for example, intonation, we are blinded from seeing that many 'problems which are central to linguistic theory' are in fact intonational - or at least dependent upon prosodic considerations (as specified by e.g. Bolinger 1961b, Crystal 1975, Halliday 1967, Parker Rhodes 1978).

Labov's assumption seems to be that it is necessary for a class of linguists to labour in the pre-theoretical gloom of some empirical vineyard until another class of linguists conceive the time ripe to apply the results of those labours to the solution (or dissolution) of theoretical problems which are raised as if they had not been considered or envisaged by the first class of linguists. Whether the assumption is based on a pietistic regard for 'specialism' (intonology), or on the arrogance of the 'theorist', is immaterial; by any sensible account of how the questing intellect works (Popper 1959; Kelly 1955; Lakatos 1976), it is nonsense. Whatever else it is based on, Labov's rejection of the use of intonational features as variables is not based on any arguments which account for the difficulty of getting those variables. [N]

[N] His silence on the matter must surely militate against an interpretation of Labov's rejection of intonational variables as being methodologically motivated. Though a motive exists, as follows.
The criteria which he uses to (positively) identify alternations between e.g. non-casual and casual 'styles' are so-called 'channel cues' which are realisations of features from prosodic and paralinguistic systems (Labov 1966). Since 'style' is one of the dimensions against which the frequency of variables is measured, the use of intonational variables would confound that dimension.

To the reasons for the difficulty in getting intonational variables I now turn. We, as ordinary hearers, and as linguists, suffer differential degrees of ignorance about the "variables" which we interpret, and the "variables" which we use. In the case of intonation our ignorance stems from two types of greasy complexity which, to make matters worse are interdependent. These are (a) the co-distribution of two kinds of intonational variability, and (b) the functional multiplexity of intonational realisations.

The two kinds of intonational variability I have in mind are what I have called realisational variability and varietal variability (Pellowe & Jones 1977, 1978a, 1979). Very many intonational features, unlike segmental phonological features, have exceedingly variable domains of realisation. For example, a feature such as allegro, in a system of tempo, may have as its domain two syllables or twenty two syllables; a tone-unit may be coextensive with one word, or with three complex clauses; a stretch of speech which carries the feature high from the system of pitch range may occur within a tone-unit or across the boundary between tone-units. The distribution, and co-occurrence, of such variations as these is what I mean by realisational variability. (This kind of variability gives rise to regulatory, or intrinsic social information, to which I return below.)

Not only do tokens from intonational systems vary realisationally, in the above sense, but as a matter of widespread naive and professional anecdotal knowledge, they vary varietally as well. That is, intonation systems exhibit regional and social differences which seem to be typologically similar to those shown by vowel systems, or consonant systems, or syntactic systems. (This varietal variability gives rise to indexical, or extrinsic social, information, to which I return below.)

It is important to stress that realisational variability and varietal variability are manifested simultaneously in any particular utterance. That is, the distinction is conceptual rather than physical. In other words, for any given segment of an utterance, there is only **one** formal co-presentation of that bundle of relevant phonic values (of duration, of frequency and of amplitude), but there are **two** types of variability patterned within it.

The multiplicity of the functions of intonation is well-expressed in the literature. It is variably believed that intonational systems express a cluster of attitudinal or affective functions (e.g. Kingdon 1959); and/or a cluster of grammatical functions (e.g. Halliday 1967); and/or a cluster of discourse functions (e.g. Pilch 1977, Brazil 1975). Much of the material has been critically compared and contrasted by Crystal (1969, 1975). More recently, Cruttenden has argued that intonation " operates with its own set of meanings which are of higher abstraction than those of grammar, attitude or discourse" (1981: 77). However, none of these stances towards the multiplicity of intonational functions quite fits the problem I am addressing, which is the distinguishability of realisational from varietal variability in one and the same formal intonational string.

I indicated above that realisational variability gives rise to regulatory, or intrinsic social, information, and that varietal variability gives rise to indexical, or extrinsic social, information. By intrinsic social information, I mean that when a speaker and a hearer are in focussed activity together, the speaker attempts to indicate her affective states, her attitude to what she is saying, and the context of its saying; to specify the roles which are to be attributed to syntactic structures, and to indicate the relative informational importance of different parts of her utterance, including those parts which are to be deemed by the hearer as having been deleted. Clearly the information which will result from the speaker's attempts in these matters cannot be at all definitely predicted by the _speaker_ in advance of the hearer's response, since the information results from the fates of the _hearer's conjectures_ rather than from the force of the _speaker's intentions_. Nevertheless, I have called this type of information _intrinsic_ because it _might_ be agreed upon by participants (cf. the kind of data generated by 'eavesdroppers' (McGregor 1983), and by people analysing themselves on film (Kreckel 1981)). The information is _social_ because it is dialogic. In projecting such information the speaker is regulating, or commenting upon, or celebrating the nature and course of the interaction, and the contributions of its participants, up to that moment. In freely inferring about the quality of such regulations, comments and celebrations, the hearer makes his contribution to the information at the dialogic interface. [N]

[N] Of course I am abbreviating the discussion here by assuming that both participants entertain a single, mutual, sense of the interface (B,C2). This is, very usually, a gross simplification (B,C3).

By extrinsic social information I mean the kinds of inferences which hearers regularly make concerning non- or extra- linguistic characteristics of the speaker. Such inferences are based in part, or in whole, on what the hearer conjectures to be characteristic general patterns of intonational realisation in the speaker's speech. What differentiates this, extrinsic, social information from intrinsic social information is that its acquisition by the hearer is not part of the speaker's focussed intention. Indeed the speaker may be completely unaware of the basis, of the nature, and even of the making, of such inferences on a given occasion. Such unknowing is often revealed in the post hoc unravelling of misunderstandings. [N]

[N] There is a class of misunderstandings which appears to arise from the hearer's having projected extrinsic social information too strongly upon the meaning or the context of the speaker's utterance. This is revealed by an unravelling statement by the hearer such as :"I thought that since you are x, that when you said z, what you meant was w ". Here x is some extrinsic social information about the speaker conjectured by the hearer. The speaker can contest the strength or the validity of x either directly ("where did you get the idea that I was x from ?") or indirectly ("well you'll have to think again, because I didn't mean w, I meant a "). Misunderstandings are extremely important data. Their absence from the literature is one more indication, were more needed, that linguistics is upside down. See Humphreys-Jones (1987).

I use the term information in a deliberately extended sense. When a hearer interprets the intonational realisations of a speaker, the social information (extrinsic &/or intrinsic) which he obtains is 'real' for him; that is, people do not seem to doubt their conjectures when they make them (which is not to deny that they may be willing to attach a probability of less than 1.0 to some conjecture). However, it may not be the same information as that derived from the same utterance by another hearer, and neither need bear any direct correspondence with the speaker's own view of the social information which inheres in her utterance. Thus, the social information in a particular _utterance_ is a potentially multi-valued function of not only the signal, but also of the source, and of the receivers. It seems impossible, _a fortiori_, that competing values of such a function could be sorted out by some truth criterion - one supplied, for example, by a 'professional' observer (B,C4).

The relationship between realisational and varietal variability on the one hand, and extrinsic and intrinsic social information on the other, is indicated in Figure B1 (Vol. III). It must be emphasised however, that it is only possible to make such gross distinctions as are illustrated there by ignoring the huge range of intentions (and their realisation) which is open to both speakers and hearers in any interactive situation.

Clearly, for example, it is open to the speaker to make varietal variability a source of _intrinsic_ social information. That is, it can be part of the speaker's focussed intention that the hearer acquires social knowledge which is relevant to (the speaker's view of what might constitute) an understanding of the current utterance (i.e. intrinsic information) not from realisational intonational differences, but from varietal intonational differences. Such behaviours may occur, for example, in jokes, in self-deprecation, for the sake of biographical veracity, in the evocation of attitudes stereotyped to that variety, and so on. Conversely, certain patterns of realisational variability in the utterance of the speaker may be conjectured by the hearer as carrying extrinsic social information instead of (or as well as) intrinsic social information. Thus [miscellaneous anecdotal data] a speaker chiding her hearer but wishing to be heard to be doing so 'pleasantly' (intrinsic), had co-occurrent realisations of low (pitch range), piano (prominence), and glissando(down) (rhythmicality), but was interpreted as having been 'too worn out to tell me off' (extrinsic).

The hearer's intention may be covert as far as the speaker's knowledge of the interaction goes, as in this last example. Clearly, however, the hearer can make his intention overt, as when he treats realisational variability extrinsically - sarcasm, attacks on the identity of the speaker, self defence, self deception. Or when he treats varietal variability intrinsically - linguistic prejudice, social distance, disdain. And both of these overt patterns can be highlighted in humorous, or quasi-humorous, frames, to which a valid speaker response is 'you're deliberately misreading me', if, that is, she does not laugh. Many kindred examples of functional-formal asymmetry in the uses of intonational systems render these two pairs of distinctions (Figure B1) less clear-cut than they may appear to be, but no less useful for all that.

To illustrate one of the motivations for the distinctions made above consider the intonation of the utterance in Figure B2.
All the tones in the utterance are level tones. Problems of interpretation for linguists, in assuring themselves that they <u>are</u> level tones, let alone assigning them particular attitudinal, syntactic or discourse functions, are barely less than those which

a visiting, southron, R.P.-speaking hearer might have in discriminating amongst various intrinsic social meanings on the one hand, and between intrinsic and extrinsic social meanings on the other.

Do the first two level tones stand in the same (syntactic) functional relationship to each other as the elements of a fall-plus-rise (compound) tone might in some non-localised variety ? Would the function of the relationship between cigarette and factory, as realised here (drop, booster, level) be equivalently realised by a fall-rise (complex) tone on cigarette in a non-localised variety ? Would the hearer be responding comprehensibly (as far as the speaker was concerned), if he responded as to an attitude of boredom on the part of the speaker, or uninterest imputed to him by the speaker ? (These two attitudes were the dominant ones elicited by the tape from hearers who themselves spoke a range of varieties other than Tyneside.)

It seems clear to me that these and a host of other questions, of a functional kind, cannot be begun to be answered until one has a detailed understanding of the varietally different systems of forms in terms of their distribution with respect to each other (B,C5). We cannot determine [N] the intrinsic social information (interpretive and interactional) which can be conveyed by particular co-occurrent intonational features from the intonational systems of some variety, until we know in some statistical detail what is the place and the significance of those features in those systems.

[N] The impossibility of this approach might be reduced if one had access to professional linguists, each of whom had native competence in one of the varieties which were of interest.

Thus, it makes no sense to ask what are the functions of level tone in Tyneside varieties vis-a-vis non-localised varieties, because such a question invests significance in the formal identity of the two. The formal identity of the two is precisely what needs to be established before such questions can make sense. Since if, for instance, one discovered that the numerical and syntagmatic distributions of level tone in some Tyneside variety were very 'different' from those in some non-localised variety, one would have little reason to imagine that they constituted the 'same' tone at all. (Cf. the problems raised by <u>cigarette factory</u>.) [N]

[N] For the moment I pass over the question which is begged by my use of 'different'.

Indeed such findings might give a picture of the correspondingly <u>depleted</u> members of the system, and hence one could have, in principle, a projection of the <u>possible</u> functional differences between the elements in the one system and in the other. Cruttenden (1981: 83-4) recognises this difficulty - of formal pseudo-identity and of formal pseudo-contrast - and he offers two plausible conjectures on the cross-varietal equivalence of falls and rises (where the latter are considered to be functioning as falls), but he does not claim to have solved the problem in general.

So far, my discussion has indicated only the broader context of why it might be difficult to get intonational variables, and I have suggested that in order to understand varietally different systems of forms, we need to establish their distributions with respect to each other. How is this to be done ? – bearing in mind that we wish to be able to compare each speaker with every other, and that we require the similarity or dissimilarity of their intonational realisations to be representable in an appropriate manner (see above Chs. 2 & 3, for appropriacy).

Consider how other 'variables' are defined, and the kinds of things which are known about them. First, recall the nature of single segmental phonological criteria (see above, Appx. A,2). Because comparability of speakers is ensured by the hierarchic relationship between an OU, its PDVs and their States, speakers can be compared directly in terms of their standardised scores on PDVs and States, or PDVs alone (Jones 1978, 1983; Pellowe et al. 1972). This kind of comparability is a reflection of the fact that from the hearer's point of view such criteria are linguistically recognisable, and hence can be isolated and talked about. [N]

N By 'linguistically recognisable' I mean that when he hears a particular State (of a particular PDV) in a lexical item, the hearer knows what it is a variant <u>**of**</u>. In addition, he can conjecture - certainly at PDV level and possibly at State level - which limited, fuzzy, but indicable part of the lexicon the same variant will occur in.

Syllabic and syntactic criteria share with segmental phonological criteria this characteristic of linguistic recognisability (see below, Appx. C). For instance when a speaker realises certain words with anaptyctic vowels, always apparently either [ə] or [ɩ] -

['fɩləm] [ɷmbəɹ'ɛlɛ] [ɹɒɩ'jalɩt͜ʔɩ] [aŋɩ'zaɩjət͜ʔɩ]

<u>film</u>, <u>umbrella</u>, <u>royalty</u>, <u>anxiety</u>

- hearers recognise what the variants are variants **of** (i.e. ± V {V= ə/ɩ}), and the forms (i.e. words) in which they are embedded. However, our capacity to predict other forms which will carry one or other of the variants is <u>less</u> than is the case for segmental phonological variants. It is likely that <u>elm</u>, <u>realm</u>, would be candidates for anaptyxis. But what about <u>loyalty</u>, <u>embrasure</u>, <u>embraces</u>, <u>alms</u>, <u>haulm</u>, <u>anxious</u>, <u>gangster</u>, to cite a few formally related extensions of the list given ? Our reduced capacity for conjecture is presumably related to our lack of any detailed understanding of what promotes a word as candidate for anaptyxis. (And this may well be a case in which the linguists' ignorance is not much more than that of hearers (B,C5b).) Similar considerations also apply to syntactic criteria such as emphatic clause enclitics,

My skirt's too short, **this**;

I could just go a toasted sandwich, **me** (see below, Appx. C). Here the conjectural problem for hearers is what lexical and syntactic constraints there are on the grammatical subject for this criterion to be able to occur. (It seems fairly clear that the subject must be a pronoun or a noun without postmodification.)

Such problems of linguistic recognisability as these are of direct relevance when we try to define criteria. When we wish to compare speakers in respect of such variants we have to establish some measure of the criterion which provides a monotonic representation of the differences amongst the speakers. (I.e. where small differences between speakers are represented by small differences of the measure.) Let us assume, in general, that such a measure would be given by a percentage on some structurally relevant denominator (B,C6). Clearly, the obviousness of what that denominator should be is proportional to our conjectural certainty about what the variants are and when they can occur (i.e. the linguistic recognisability of the variants). Nevertheless, the point for the present argument is that though we may not know enough about such variants to be able to establish an optimal denominator, we can always specify one which will be statistically and structurally related to it, though not necessarily in ways which are exactly specifiable. [N]

[N] An optimal denominator, given the specification of the variants, will permit the measure to range from its minimum to its maximum in its application to a representative sample of speakers. E.g. if we define the denominator for anaptyctic vowels as 'the total number of words' the range of the measure will be smaller than if we define it as 'the total number of words having consonant clusters'. But if we assume, as we reasonably may, that the maximum of anaptyctic vowels for any speaker **must** always be less than the total number of that speaker's words with clusters, then neither of these measures would approach its maximum.

Intonational variants differ from others in not being linguistically recognisable in the sense in which I have been using that phrase. In the case of the other variants there is a class of forms and/or environments (albeit about which our knowledge might not be very precise) in which two or more variants are able to be realised. These variants are the states of a criterion which is named according to the form (or environment) and the nature of the variants. But in intonation the variants

are not associated either with a class of forms or with a class of environments such that they can be recognised. Thus, <u>intonationally,</u> we cannot directly perceive variants as between (a) $_{NL}$ [gɷd dL <u>lɔd</u>] (drop+level), and

(b) $_{LT}$ [ɑ dɩvn̩ dL <u>naː</u>] (drop+level)

since both can have exactly the same <u>drop+level</u> range, prominence, length, and so on. On the intonational evidence alone here, we cannot tell whether what (a) and (b) are saying does or does not distinguish them. Of course we 'know' that real hearers treat clusters of variables in either formatives or sentences holistically, so that in this example, for them, this problem almost certainly does not arise. But the example given here is a very marked and clear case, and less marked cases might provide real hearers with a problem akin to the one I am addressing. In addition we have to beware of not perpetrating another correspondence fallacy (Bazell 1966, and see above, Chs. 2 & 3, & passim).

We do **not**, then, suppose that there is a form called falling tone which has, say, two realisations, one wide, one narrow, and that speakers with **x** vs. z non-linguistic characteristics would be correlatively predisposed to use wide vs. narrow realisations (B,C7). (This would be the parallel of the case of [ei] vs. [iə] in great.) The impossibility of defining intonational criteria in this way is a reflection of the fact that not only are the functional ends of intonation shared by all 'normally speaking' members of the community, as far as we know, but so are its realisational means. [N]

N Perhaps it is necessary to consider the possibility that different subsets of the population **do** have non-identical sets of functional ends in respect of prosodic-paralinguistic expression. This would certainly avoid the problem of having to argue that terminal rises in declaratives in Belfast are "in some sense falls" (cf. Cruttenden 1986, and Nolan's review of same (Nolan 1987)). But at present we cannot even **imagine** being able to imagine what techniques could establish such differences of functional end.

The problem then, is that in respect of realisational means there are not even apparent discontinuities between linguistically different groups of the community in respect of their use of the terms in the various prosodic systems specified by Crystal. 45

It might be thought that we could capture variability in respect of tones, for example, by reference to the tone units in which they occurred (B,C8). We might specify some such criterion as 'simple (non-modal) finite positive declarative clause' and have as its states all of the tones. Figure B3(a,b,c,d) shows how such a criterion might be used to establish similarity between four imaginary speakers **g, h, i, j**. Such an approach looks promising, even if we have no idea whether similarity of type I or type II (Figure B3 c,d) is the best representation of the matter. N

N Strict binarisation (as represented by type IV similarity) is clearly far too strong a reduction technique for systems about whose variability we know so little. I return to this problem below in Appendix D.

There are however, four major objections to such a method of defining intonational criteria:

(i) it can take no account of the narrowing or widening of the domain of the variant which would follow from marked tonality;

(ii) there are reasons to suppose that marked tonicity will affect the capacity of the criterion to be monotonic with respect to all speakers (see below, Appx. B, section E);

(iii) it is known that within a variety a particular syntactic form can be realised with different tones for different kinds of affective, pragmatic and thematic function (Halliday 1967) [N] ;

N That is, the criterion, as defined in the last ¶, would be as likely to elicit the realisational variability as the varietal variability of a particular speaker.

(iv) the relevance of either a sense group, or of a particular kind of syntactic structure, to the definition of criteria from 'less linguistic' systems (Crystal & Quirk 1964), is even less apparent. Therefore such forms of definition would separate - before the fact - systems about whose variant behaviour nothing is known.

For all the reasons discussed above, it seems that we are unable to define intonational criteria in terms of the structure, the interdependencies, or the functions of the intonational systems whose variability they are to represent. Since the criteria have to be capable of distinguishing between localised and non-localised varieties, one possible solution to the problem of defining criteria is to determine the extent to which facts known about the distributions and functions of prosodic and paralinguistic features in non-localised varieties fit with those known about localised varieties. I report a pilot effort in this direction (Pellowe 1970d). [N]

[N] As far as I know, before this pilot work was done, there had been no investigation of the intonational facts of Tyneside, or indeed of North Eastern, varieties. I.e. there were no known facts about the intonation of localised varieties. But later, for e.g. Liverpool and Belfast see Knowles (1978) and Cruttenden (1981), respectively.

B (i): Getting criteria empirically.

Methods used by Quirk et al.(1964) to establish the correspondence between grammatical and prosodic features in the speech of 'educated people' (Quirk 1960) were applied to a small corpus of localised Tyneside speech. Comparison between the two sets of results led to the specification of some prosodic criteria which were thought to be promising for classifying a representative range of speech varieties – from non-localised to localised.

The localised material consisted of utterances from two speakers: (a) male, 35, in conversation with myself, (b) female, 55-60, monologuing (with two hearers/eavesdroppers, who did not sound (and did not themselves feel) as if they were addressees). [N]

[N] I am grateful to John Sword for making this recording and for giving me a copy of it.

Both recordings were made surreptitiously. I refer to the man as K and the woman as J. The analysis is of the systems of tone and of pitch range (Crystal & Quirk 1964) and their co-occurrence patterns with certain grammatical features.

Crystal and Quirk (1964) set up a tone system of seven terms - fall (F), rise (R), level (L), fall-rise (FR), rise-fall (RF), fall-plus-rise (FpR), rise-plus-fall (RpF). This is a singulary system. (This is unlike Halliday's (1967) system, in which there is a primary system at tonic, equivalent to the selection of one of the above tones; and secondary systems at both pretonic and tonic). Instead of secondary systems of <u>tone</u>, Crystal and Quirk set up a system which is potentially independent of tone, that of pitch range. There are two subsystems (simple and complex) relating respectively to non-kinetic syllable-long pitch changes, and kinetic pitch changes in the polysyllabic segment. Simply regarded, tones <u>must</u> be kinetic pitch changes (or, in the case of level tone, a pitch sustention), whilst all other contrasts may be described by reference to the system of pitch range.

The analytical framework and rationale of Crystal & Quirk (1964) is adopted because there is no other method of analysis of these phenomena which is as flexible, as relevant or as delicate as theirs. In addition, there is a good deal of material analysed using their systems, and so a useful volume of material exists for comparative study (cf. in particular Crystal & Davy 1969).

Experimental studies by Bolinger (1965) and Lieberman (1967), show

(1) that hearers perceive pitch not as a series of different levels, but as configurations or changes in direction;

(2) that the fundamental importance of relativity, in the analysis of intonation by (phonemic) levels, leads to serious inconsistencies of analysis;

(3) that the linguistic importance of intonation from the hearer's point of view is not reproducible from a machine analysis of duration, frequency and amplitude.

The polysystemic approach of Crystal & Quirk effectively permits the examination of the variability and the modality of distributions amongst various systems. The relationships between individual terms in different systems can be sought in scalar terms, and the stylistic norms of correlation between terms or systems can be determined for individuals or groups.

The general approach of Crystal & Quirk (1964) has also two overt methodological advantages for the types of methods proposed in this research:

(1) The different systems can be examined with varying degrees of fineness or delicacy, and this is of extreme importance in establishing varieties from a data base. Where it is found necessary or desirable to treat a system with more delicacy at a later date , there will be no loss of generality and no need to re-organise coding methods or rewrite criteria;

(2) the systems and their notation were developed in connection with the Survey of (Educated) English Usage, and while it is by no means possible to say that the varieties used by their informants are always non-localised, certainly a fair number must be. In this sense, the correlational norms established for between- and within- system variance (Crystal 1966, 1969) have direct relevance for the non-localised and some of the transitional-localised varieties under consideration here.

Quirk et al. (1964), in their investigation of grammatico-prosodic correspondences, examined a corpus of c. 10,000 words, yielding 1880 tone units (both super- and sub- ordinate tone units (TUs)). Whereas the corpus of localised speech used in this pilot investigation was of c. 1500 words which yield 245 TUs. The overall average number of words/TU in the study of Quirk et al. was 5.3, and was 6.1 in this localised corpus (L-corpus). Figure B4 gives the % distribution of TU lengths (length in words). The trace for the data of Quirk et al. (hereafter referred to as SEU-corpus) is somewhat hypothetical, since it is inferred from only three points and the arithmetic mean, which is slightly lower than it is for the L-corpus. The degree of disparity between the distributions is not as significant as it looks, and only for TU word-lengths 9 to 12, if at all. The modal average (point of maximum distribution) is the same for both corpora, that is 4 wds per TU for 15% to 17% of TUs. Length of TU then, will **not** serve as a differentiating criterion as between localised and non-localised varieties.

The next most simple thing to look at is the frequency of occurrence of different tones. These are given in percentage form in Figure B5. Each corpus has a different column. The two speakers in the localised corpus are also given separately, where their percentages are recalculated. One may observe that:

(a) The 27 % points lead of Fs over Rs in the SEU-corpus is reduced to a 17 % points lead of Fs over Rs in the L-corpus;

(b) In the SEU-corpus, Ls (2%) are relegated to a minor category, but in the L-corpus Ls (14%) form the third biggest category. Looking more closely, we see that the increase of Rs and Ls in the L-corpus seems to be at the expense of Fs;

(c) But if we look at the figures of tone frequency broken down for K and J individually, we see that, relatively speaking, both Ls and Rs are about half as frequent as Fs for K, but that for J the distribution is very similar indeed to that of SEU-corpus.

These observations raise two points:

(i) a simple tone frequency count might provide a criterion which would show the macro-differences between maximally localised and typically non-localised speakers in terms of intonation; but

(ii) it is necessary to find further criteria which will

(a) show the differences between J and SEU-corpus, and

(b) show the affinities between J and K. [N]

[N] But given my discussion of Bazell's attack upon linguistical fallacies and his demonstration of the unprovenness of systemic solidarity, there is considerable doubt as to why we should assume <u>any</u> kind of principle (B,C9) which - in two varieties which show overall similarity - expects <u>all</u> the subsystems of those two varieties to be similar to each other in roughly equal degree.

In terms of further comparisons between the L-corpus and the SEU-corpus, these two requirements may be be amplified as follows:

(1) we would like to know if the grammatical variables which co-occur with certain tone and pitch-range system sequences are different as between localised and non-localised varieties;

(2) we would like to know if the modal distributions of terms from the pitch-range system with terms from the tone system differ as between localised and non-localised varieties;

(3) we would like to know the nature and extent of different polarities and conditionings in localised intonation, which would

lead to different replacement norms (Quirk and Crystal 1966) from those of non-localised varieties (see below, Appx. B, section E).

In that pilot study I examined only the simplest grammatico-prosodic co-occurrences and tonic-pitch range co-occurrences (i.e. (1) & (2) above).

Apparently stable differences between the corpora in terms of these co-occurrences were taken to indicate criteria which would be useful for a classification of speech varieties (B,C10). In respect of the SEU-corpus, I only examined the intersection of nucleus-type (F or R etc.) with the form-class of the item carrying the nuclear syllable (noun, adverb, etc.).

Because it was only relevant to the purpose of Quirk et al. to discuss major categories (unfortunately thereby losing Ls) these comparisons are restricted to the following terms:

<u>form-classes</u>	noun (common)	n.
	noun (Name)	N.
	adverb	adv.
	premodifying adjective in nom. gp.	p.adj.
	adjective not operating as modifier (free)	f.adj.
	pronoun as Subj. in	

	clause structure	pron.
tones	fall	F
	rise	R
	fall-rise	FR
	rise-fall	RF
	fall of fall-plus-rise	F(FpR)

First of all, in Figures B6a & B6b, I plotted the percentage at which any given tone occurred on the given form-classes (where 100% refers to these form-classes only). The plots for each tone are given separately - the continuous line is the SEU-corpus (standardised), the broken line is the L-corpus (standardised). The following points emerge:

(1) Looking at the distributions of F and R (the first two blocks of Fig. B6a) we note that L-corpus :

(a) has a high relative % of F but a low relative % of R on common nouns (n);

(b) has a high relative % of R on adv.s and pron.s (The converse of these and subsequent remarks should be taken to apply to SEU-corpus, unless otherwise specified.)

(2) Looking at FR and RF (Fig. B6a block 3 and Fig. B6b block 1), L-corpus:

(a) has a low relative % of FRs, but a high relative % of RFs, on n.s ;

(b) has %s of FRs and RFs on N which are, relatively, equivalent to those in SEU-corpus;

(c) has a medium high relative % of FRs on adv.s and pron.s

(3) Looking at F(FpR) (Fig. B6b block 2), L-corpus has high and medium high %s of F(FpR) on, respectively, n.s and pron.s

Figures B7a & B7b present the same material transposed to a form-class base (i.e. with scores re-standardised). That is, we may now read off, say, the percentage of (nuclear-syllable) common nouns which carry F as against R etc.

The following points suggest themselves:

(1) the distribution of nuclear common nouns is much more even across all the itemised tones in SEU-corpus than it is in L-corpus (where 70% of nuclear n.s carry F);

(2) the low relative % of nuclear adv.s carrying F in L-corpus is accounted for by a high relative % of nuclear adv.s carrying R and

FR. [N]

[N] Many of the distinctions so far picked out may be conjectured to provide secondary evidence for a different pattern of localised replacement norms (B,C11). By replacement norms are meant those groups or pairs of tones which can tolerate bi- or uni- directional exchange (Quirk & Crystal 1966). In particular, for Names, we see that both for L-corpus and for SEU-corpus there may be an equivalence (though of different frequency relations) between FR and RF. In general as the tonic selections for a particular form-class become less numerically likely, they may become more grammatically equivalent or, at least, less grammatically distinctive.

---

(3) in L-corpus a high relative % of padj. carry F, and a high relative % carry R, but a low relative % carry F of F(pR);

(4) Apparently as a corollary of (3), L-corpus has a very low relative % of fadj. carrying simple F, but a very high relative % of fadj. carrying complex F (of FpR).

Though the comparisons presented thus far by no means exhaust the complexities of the material and the analyses of Quirk et al. (1964), I next considered differences of co-distribution between terms from the systems of pitch-range and tone: in particular the modal (i.e. maximum) co-distributions.

We have already obtained simple criteria from the system of tone which are capable of distinguishing between localised and non-localised varieties. However, those criteria of tone frequency were too crude on their own, because they put J too close to the SEU-corpus and too far away from K. One must therefore suppose that there are other criteria which will indicate the distance of J from the SEU-corpus and her proximity to K. I hoped to find these in the pitch-range system , or at least in the way the pitch-range system is realised in conjunction with the system of tone.

The pitch-range system (simple) is described in detail by Crystal and Quirk (1964). Essentially, and informally, one may say that a non-kinetic change of pitch can affect either nuclear or non-nuclear stressed syllables in one of seven ways: by means of a low drop (ld), a drop (d), zero (Ø), a continuance (c), a booster (b), a high booster (hb), or an extra high booster (ehb). (Criteria for their recognition are given in Crystal & Quirk (1964).)

To investigate, crudely, the co-occurrences of terms from these two systems in localised and non-localised varieties, I compared published analysed texts of SEU (Quirk et al. 1964) with the L-corpus. (I realise that these published texts may not be entirely representative, though how "representativeness" might be established in this case is itself a difficult matter.)

The total number of TUs being compared is:

		Tones		
	TUs	L	R	F
L-corpus	193	37	50	106
SEU-text	127	6	35	86

The % distribution of tones in SEU-text compares reasonably with that in SEU-corpus. (It might be argued that some of the distinguishing differences discussed below arise from analyst differences, that is from different interpretations, or applications, of the criteria given in Crystal & Quirk (1964). I deal with this eventuality all too briefly below, Appx. B, section F.)

Two kinds of pitch range system phenomena were examined in the corpora (SEU-text; L-corpus) depending on their place in the TU under scrutiny. Each TU was scored depending on:

(a) whether a term (drop, booster, etc.) was used <u>on</u> the nuclear syllable - and if so which term it was;

(b) whether a term (drop, booster, etc.) was used anywhere <u>before</u> the nuclear syllable in the same TU at whatever distance - and if so, which term it was.

Figure B8 gives a table of the analysed combinations of (a) and of (b) which came up in the two corpora. Thus row 1 under L (first entry) signifies that (in raw scores) 4 Ls (out of 6) in SEU-text had no pitch range system term on the nuclear syllable, and no pitch range system term in the preceding remainder of the TU. Jumping three entries in row 1, we see that none of the 36 Ls in L-corpus came into this category, and so on.

Figure B9 gives the same information in % form. Figure B10 condenses the material of Figure B9, and groups terms of the pitch range system affecting the nuclear syllable, regardless of the pitch range terms in the rest of the TU. This gives a distribution of Fs, Rs, and Ls on the basis of

(a) zero pitch range modulation (Ø),

(b) pitch range modulation by a 'generalised drop' (i.e. by either drop or low drop),

(c) pitch range modulation by a 'generalised booster' (i.e. by one of continuance, booster, high booster, extra high booster). The distributions are each % for SEU-text, J (who together with K makes up L-corpus), L-corpus, and K, in that order. (The figures are ordered this way to shew the trends which I now discuss.)

Here the required proximity between J and K (see above p. 170) is beginning to be expressed. We can think of these trends as putative criteria for distinguishing localised from non-localised varieties (B,C12). Nonetheless it is possible to rank the trends in terms of our confidence in their likelihood of doing this. This I now do, ranking them from greatest to least likelihood.

One might want to claim that localised varieties may be distinguished from non-localised by having:

(1) more Ls with boosters,
(2) fewer unmodulated Ls [N],
(3) more Rs with drops,
(4) fewer unmodulated Rs,
(5) fewer unmodulated Fs,
(6) more Fs with boosters,

(7) more Ls with drops,
(8) more Rs with boosters,
(9) more Fs with drops.

[N] This trend is ranked here on the basis of the face-value of the figures (which are small). However this trend is supported by that for Ls and boosters. We may group the ranks, in terms of the strength of trends they shew, in four bands: (a) 1 & 2; (b) 3 & 4; (c) 5, 6 & 7 (d) 7, 8, & 9.

In spite of the different strengths of these trends in their conjectured capacity to distinguish localised from non-localised varieties, it is worth pointing out that for all these tones, localised varieties favour co-occurrence with any marked pitch range term compared to non-localised. (And conversely for tones with zero pitch range modulation.)

Pellowe (1970d) speculated that there was an inverse relation between the gradience of the trends in Figure B10 and their stable discriminatory power as between localised and non-localised varieties. (Gradience increases as one moves from (1) to (9) in the list above.That is, gradience is the inverse of contrastivity (Bolinger 1961b).) He went on to conjecture that an interesting test of such a principle of gradience amongst co-occurrences of tone and pitch range would be in transitional localised varieties.

Such varieties seem to arise as two quite distinct types:

Type <u>**a**</u>: a non- or other- localised speaker may adopt a widening range of localised features; [N]

[N] This is not necessarily equivalent to accommodation in the sense of Giles (e.g. Giles and Powesland 1975), since the realisational habit may persist. Such a possibility seems not to have been recognised by accommodation theorists.

Type <u>**b**</u>, a localised speaker may substitute for diagnostic localised features some approximation to the non-localised features which are construed to be equivalent by that speaker.

We should find that for co-occurrences where the gradience is strongest (i.e. most continuous), there is <u>least likelihood of change</u> in either kind of transitional variety. Under this conjecture a transitional-variety speaker of type <u>**a**</u> will adopt an increased relative frequency of BR less quickly (if at all) than a vastly increased relative frequency of BL. (This is confirmed by informal observation.) Conversely, a transitional variety speaker of type <u>**b**</u> will increase the relative frequency of ØF and decrease the relative frequency of BF before decreasing the relative frequency of BR. I will examine these conjectures in the light of further evidence below (Appx. B, sections C & D).[N]

[N] According to these remarks, J herself may be interpreted as a transitional variety speaker of the second type. If this were true we would expect application of these measures to a bigger localised sample to move the LOC figures in Figure B10 closer to those of K. For a partial investigation of this see below, section D.

Pellowe (1970d), on the basis of the above analysis, proposed twelve quantitative criteria for the purpose of grouping a mixed sample of localised and non-localised speakers on the basis of grammatico-prosodic resemblances between them. They were:

(1) the relative standardised frequencies of F:R, F:L, and R:L;
(2) the absolute standardised frequencies of all tones;
(3) the relative %s of Fs and Rs on common nouns;
(4) the relative %s of Rs on adv.s and pron.s on the one hand, and on common nouns on the other;
(5) the relative %s of FRs and RFs (on both common nouns and Names);
(6) distribution of all nuclear common nouns across all tones;
(7) relative %s of nuclear adv.s with F vs. those with R & FR;
(8) relative %s of nuclear padj.s with F and with R;
(9) relative and absolute %s of Fs with B at nucleus;
(10) relative % of ØRs;
(11) relative and absolute %s of ØLs;
(12) relative and absolute %s of BLs.

Some important regularities in Fig. B10 were apparently ignored by Pellowe (1970d). Consider the figures in pictorial form (Figure B11). First a negative point: the most gradient co-occurrences, in the above sense, i.e. DF, BR, DL, may be low frequency gradients for physiological reasons of one kind or another:

(i) by definition, BR and DF will tend towards the limit of a speaker's voice range; their use might therefore be expected to be associated with very marked emotion (e.g. hysteria and grief, respectively);

(ii) by its nature, L may be less salient for hearers than the kinetic tones (for evidence possibly favouring this view see Quirk & Crystal 1966, and below, section E). Were this the case, its co-occurrence with drop, in anything other than utterance final position, might simply be construed as the occurrence of drop alone. If this were the case for hearers, and if speakers were aware of their behaviour as hearers in this respect, then speakers might avoid the pattern (and cf. Bladon (1986) for support of such a notion in other respects). [N]

[N] This interpretation is rather weaker than (i), immediately above, partly because the gradient itself is weaker, and partly because perfectly plausible examples come to mind:
*I /want you to /F+go..[/dLnow]..if you dont /+Rmind.
However, I guess that in a majority of cases there would be some indicative co-occurrent feature of reduced pace (pause, drawl, lento, rallentando).

Secondly, Figure B11 reveals an interesting mirror image relationship between DR and ØR on the one hand, and BF and ØF on the other:

(i) the relative frequency of the members of each pitch-range/tone pair is reversed as between localised and non-localised varieties;

(ii) a small divergence in the relative %s of ØR and DR is indicative of non-localised speech; a wide divergence of localised;

(iii) the relationship between ØL and BL is similar to that between ØF and BF, though the figures are less trustworthy.

Such additional observations indicate that stable criteria for distinguishing between varieties according to intonational differences might, in addition to those above, be defined in terms of the <u>relations</u> between different co-occurrent pairs of terms from the systems of tone and pitch range (B,C13). The feasibility of this is examined below (in sections C & D), but here it is necessary to discuss independent, corroborative material for the <u>kinds</u> of criteria set out above (p. 187 , above). The material confirms the suitability of some of the criteria defined above, casts doubt on others, and introduces new ones.

McNeany (1971) analysed the co-occurrences of grammatical, prosodic and paralinguistic features in conversations between himself and two members of the Tyneside speech community, and between myself and a third member (i.e. 'N.L.') of the speech community. [N]

[N] It is likely therefore that the material analysed by McNeany contains less variability caused by irrelevant differences of interactive setting and participants, than is likely to be contained in the material analysed by Pellowe (1970d).

Two of the speakers who provided McNeany's data spoke localised varieties (one a man, 'L.1.', one a woman, 'L.2.') and one a non-localised variety (a woman, 'N.L.').

The numbers of TUs analysed were:

'L.1.'	'L.2.'	'N.L.'
324	284	307

The overall distribution of individual tones (%TUs) was broadly similar to that found by Pellowe (1970d), (cf. Fig. B5.) However there are differences of level of distribution. In particular:

(i) 'N.L.' has 70% F compared to SEU-corpus 51% F, and this is directly compensated for in the distribution of rises - 'N.L.' 4% R, SEU 24% R;

(ii) neither 'L.1.' nor 'L.2.' have such high relative %s of levels (5% & 7% respectively) as K (19%).

Nevertheless, McNeany's data support criteria (1) and (2) above (p. 187 above, see also Figure B12 below).

For criteria concerning the interaction between tone and form-class, there are no data in McNeany (1971). However, his analysis of the interaction between tone and pitch range casts doubt on some of the trends observed above as follows (see figs. B10 and B11):

Co-occurrences	Trend (from non-localised to localised in Pellowe (1970d))	Trend (in McNeany (1971))
ØL	% decreasing widely	same
BL	% increasing widely	same
DL	% increasing	same
ØF	% decreasing	weakly opposite
BF	% increasing	same
DF	% increasing slightly	equivocal
ØR	% decreasing widely	strongly opposite
BR	% increasing slightly	same
DR	% increasing widely	strongly opposite

We might wish to attribute these reversals to one, or all, of the following:

(a) insufficient numbers of TUs in Pellowe's (1970d) data,

(b) the differences of setting and participants referred to above,

(c) differences between the auditory and cognitive habits of Pellowe and of McNeany.

Doubtless all of these contribute something to the problem, but the figures indicate that McNeany's informant 'N.L.' is the source of most of the reversal of trends. If we compare 'N.L.' with SEU-text, and 'L.1.,L.2.' with LOC, in respect of the troublesome trends (ØR, DR, ØF), we find:

	'N.L.' compared to SEU-text has	'L.1.,L.2.' compared to LOC have
ØR:	41% fewer	18% more
DR:	51% more [N]	15% fewer
ØF:	16% fewer	5% more

[N] Inspection of the transcript does not reveal anything odd, or marked, about the syntactic distribution of DR. It occurs in subordinate clause TUs (both conditionals and relatives), you know tags, heads of subjects and complements, and lexical and modal verbs.

I am unable to see any necessary connection between these facts about 'N.L.' and the fact, cited above, that she had 20% more falls than SEU-corpus. The only conjecture I make – an extremely weak one based on what I recall of the conversation – is that 'N.L.''s conversational manner is idiosyncratically affective (B,C14). But I do not know of any research which has indicated that such a manner might be associated with DR rather than ØR. [N]

[N] Could it be that DR, in the mouth of 'N.L.', is equivalent to $_{LT}$ BL; i.e. that this is her attempt at being localised ???

The problem can only be resolved by the inspection of further material: especially if the direction of the trend can be established from varieties which are differentially localised - a set ranging from maximally to transitionally localised.

McNeany (1971) proposed ten criteria in addition to those listed above, for nine of which there are no comparable data in Pellowe (1970d). The one criterion for which there is data from Pellowe - % Fs with drop at the nuclear syllable - is not corroborated by it (cf. the trace for DF in Figure B11). A second criterion - % R co-occurring with wide (in the complex system of pitch range) - does not appear to be supported by McNeaney's own data: 'L.1.': 1.6%; 'L.2.': 6.1%; 'N.L.': 7.7% (and the raw figures are very small). His other suggestions seem to me important, and his distributional findings are indicative of criteria likely to have good discriminatory power. I quote relevant figures from pp.1-3 of a table of "Grammatical (sentence-type) co-occurrences" (McNeany 1971):

	'L.1.'	'L.1.'	'N.L.'
(1) Falls on statements, %statements	59	59	82
(2) Generalised falls on statements, % statements	68	62	87[N]

[N] 'Generalised falls' comprise the sum of falls, rise-falls and rise-plus-falls.

(3) Falls on questions, % questions	26	56	76
(4) 'Generalised falls' on questions, % questions	39	58	92
(5) Levels on interjections, % interjections	18	34	2
(6) Interjections with any non-simple tone, % interjections	10	13	32
(7) Wh-questions with R, % wh-questions	71	100	0
(8) Polar questions with R, % polar questions	10	7	50

Figure B12 summarises 22 prosodico-grammatical criteria for classifying a range of speech varieties ranging from (Tyneside) localised to non-localised. Notice that criteria 1-8 are compound criteria involving, as they do, more than one measure.

In the case of each of these criteria it is important to emphasize possible limitations on their representativeness of the Tyneside speech community at large. The possible limitations arise from five characteristics of the data from which the criteria were derived, which we may consider in tabular form.

	Criteria (1-8)	(9-12)	(13-22)
No. and gross variety of speakers	(a) 2 L (b) 10 NL	as (1-8) ? as (1-8)	2 L (≠(1-8)) 1 NL (≠(1-8))
Sex	(a) 1 F, 1 M (b) 10 M	as (1-8) ? as (1-8)	1 F, 1 M (≠(1-8)) 1 F
Situation	(a,i) monologue with others (not JP)	as (1-8)	conversation (VM)
	(a,ii) conversation with JP	as (1-8)	conversation (VM)
	(b) panel discussion (radio)	?as (1-8)	conversation (JP)
No. of TUs	(a) 245 (b) 1880	193 127	608 307
Analysts	(a) JP (b) SEU	as (1-8) as (1-8)	VM VM

First, the number of speakers is very small. Since we have no idea what the range of prosodic and paralinguistic variation in a given statistical or physical population <u>is</u>, we must not be too confident that criteria based on such small numbers will be able to <u>represent</u> that range of variation adequately.

Secondly, what little is known or conjectured about gender differences and speakers' habitual realisation of prosodic and paralinguistic features (e.g. Crystal 1969; Guy et al. 1986) warns us that we will probably not be able to tell which of these criteria represent the male/female aspect of varietal variability and which the localised/non-localised aspect.

Thirdly, there are considerable situational/interactional differences between (a,i), (a,ii) and (b) for criteria (1-12). Since such differences will tend to be associated with realisational variability rather than varietal variability (see the discussion of Fig. B1 above, section A ; and e.g. Crystal & Davy 1969), we shall have cause to inspect carefully the effects of these criteria on situationally homogeneous speech samples. [N]

[N] Note however that the McNeany (1971) data, which are more situationally homogeneous, do support criteria (1) & (2), and therefore, by implication, might also be taken to weakly support (3) to (12).

Fourthly, the numbers of TUs analysed are rather small, especially for criteria (9) - (12), the bases of which may well be sensitive to matters of frequency.

Finally, it is predicted by the model proposed above (Chs. 2 & 3) that different linguists (as hearers) will behave differently, transcribe differently, conjecture differently, in the face of some piece of language, just as 'ordinary' hearers do. If this is true of segmental phonology (Ladefoged 1960; Pellowe et al. 1972b, and see Appx. A.6, above), it is likely to be at least as true of prosodic and paralinguistic phonology. The extent to which these dimensions may be representations not only of the various data which I have discussed, but also of the analysts themselves (cf. Ringaard (1965) and Appendix A above) is an interesting and difficult problem which I examine in part below (in section F).

In spite of all such gloomy prognoses, these criteria are clearly characterising **something** of the prosodic and paralinguistic realisations of the speakers which satisfies our definition of 'criterion'. (A criterion is any feature of speech which both shows at least two variants across the population under consideration, and is not **logically** predetermined by the nature of any other criterion in the set of criteria currently in use.) Further attempts to say what this 'something' is will be made below (in sections C & D).

Meanwhile, it is necessary to reflect on the assumption we are making in constructing these criteria, and on its interaction with the low linguistic recognisability of prosodic and paralinguistic criteria. The assumption (B,C15) is:

given the prosodic and paralinguistic systems as defined (Crystal & Quirk 1964), and the regularities of realisational co-occurrence between features of them (Crystal 1966, 1969), and between prosodic and grammatical features (Quirk et al. 1964) in educated British English,

then marked differences of the frequency or of the co-occurrence distributions of any feature or features in Tyneside English will be the basis for a satisfactory criterion (i.e. dimension of the VSp).

Aside from the gloomy remarks above, is there anything <u>intrinsically</u> wrong with this assumption ?

Notice that, unlike segmental phonological criteria, these prosodic criteria are defined in an inherently quantitative fashion. It would not make nonsense of segmental criteria, though we would lose information to be sure, to treat them as qualitative binary criteria (presence/absence). But we just do not know enough about prosodic &paralinguistic systems to be able to treat prosodic & paralinguistic criteria like that. Anyway, as I have suggested above, we probably **cannot** ever know what it would be necessary to know to do this. Two problems arise in connection with such inherently quantitative criteria as these.

First, the dangers of defining logically dependent criteria increase. There is, in fact, one clear case in Figure B12, viz. "overall % of all tones". Clearly, given seven tones (F, R, L, FR, RF, FpR, RpF), the overall percentages for any six, in respect of some speaker, totally predict that for the seventh. Less clear is whether criterion (15) of Fig. B12 is logically dependent on criterion (16) - (though clearly (15) is 'contained in' (16)) - knowing a value for either of these criteria for some speaker does not seem to enable us to predict a value for the other. Nevertheless the 'contained in' relationship here, though not in conflict with the definition of 'criterion', will introduce a certain amount of spurious similarity between speakers who have similar values for both. In the present state of our understanding of these matters this is a price worth paying in order to find out more about the ways in which the criteria behave.

Secondly, Bateson (1979b) warns, in the context of biology and genetics but not irrelevantly for the present venture, that quantity does not determine pattern:

> "It is impossible, in principle, to explain any pattern by invoking a single quantity. But note that a ratio between two quantities is already the beginning of a pattern . . . What appears to be a genesis of pattern by quantity arises where the pattern was latent before the quantity had impact on the system." (1979b: 63-4, emphasis in original).

For our present purpose I think we should take this to imply that criteria such as (1) in Fig. B 12 - "relative %s of F:R; F:L; R:L" - are to be much higher-valued than those defined on single features, such as (2), "overall %s of tones". In other words we should always try to ensure that quantitative criteria are relational, whilst at the same time guarding against spurious similarities of the type mentioned above - provided that no information can be gleaned from them.

In summary, then, there does not seem to be anything intrinsically wrong with the assumption (B,C15) that different co-occurrence frequencies will provide useable criteria -- subject to these caveats.

These considerations about quantitative criteria for prosodic and paralinguistic features are connected also with the problem of the low recognisability of variants of those features. I suggest that one of the consequences (and one of the defining characteristics) of this low recognisability is the poverty of folk-linguistic terminology for varietal variability in intonation (B,C16). Apart from a few expressions to indicate whole-system differences (usually some musical or quasi-musical term with a category label e.g. 'musical accent', 'attractive tone of voice', 'sing-song intonation', 'monotonous delivery' . . .), or attitudinal interpretations of them ('bloody-minded accent', 'sarcastic southern voice' (cf. Crystal 1969; Cruttenden 1981)), there appear to be no ways of referring to features of, or co-occurrences between, features in the more linguistic prosodic systems (sensu Crystal & Quirk 1964). In particular there seem to be no folk-linguistic ways of referring to tone (and the abstractions of tone unit, tonality, tonicity), nor to pitch range, nor, to a lesser extent, to tempo. [N]

[N] Even without considering the distinction between realisational and varietal variability, the terms for individual features in these systems show this to be the case. Comparatively, the terms used by Crystal & Quirk (1964) for features in less linguistic systems are much more closely related to (or derivable from) ordinary folk-linguistic terms.

(One might speculate that a reason for the poverty of folk-linguistic terms for items in the more prosodic systems could be something to do with their centrality to the expression of her uniqueness (rather than her group membership) by the individual speaker; that is, in terms of the moment by moment expression of identity-cum-affect. (For further discussion see Sacks (1986b: 76).)

Such a speculation must be based on the assumption that folk-linguistic terms exist solely because many persons in that culture have metalinguistic things which they wish to discuss.)

This lack of folk-linguistic terminology for intonational variability of the varietal kind is the main reason why our intuitions in the matter of evaluating competing intonational criteria are rather weak or uncertain. [N]

[N] As far as I know, no work has been done on the relation between 'professional intuitions' in some particular field and the strength and richness of folk-linguistic judgements and terminology (i.e. 'naive' intuitions). In set-theoretic terms it would indeed be peculiar were there no connection !

In fact our intuitions are so weak in this respect that there are problems - in respect both of which criteria should be included, and of their definition - which can only be solved by lengthy trial and error methods. The inclusion problem can be very simply illustrated as follows. In Fig. B4 the % distibution of TU lengths in words was found to be very similar as between 'educated' speech and localised Tyneside speech. My expectation, in bothering to depict the relationship for the Tyneside speakers and compare it with that for the educated speakers, was that there would be a difference which could be used as a dimension (criterion). But when none appeared, the possibility of using the relationship as a dimension was **simply rejected without surprise or reluctance**, rather than, for example, questioning the representativeness of the data upon which the figure was based. A weak expectation indeed (B,C17) !

But, given the smallness of the data base and our uncertainty about what criteria might usefully discriminate amongst different sub-groups of the speech community, maybe we should include as many as we can imagine of those criteria which are definable and of which the features can reasonably be expected to occur in a forty minute interview. In other words, with a different set of informants maybe '% distribution of TU length in words' would discriminate subgroups of those informants.

The definition of criteria is a problem arising from similar considerations. I have already suggested that each intonational criterion should express an interaction between two (or more) features. (Cf. Bateson's remark quoted above.) Thus there may be reason to believe that the '% difference between F and R on common nouns', or the '% difference between R and RF on adverbs' would divide up some sample of speakers in useful and interesting ways. But how can we be sure that these are the optimal ways of expressing the interaction between these variants in respect of any given sample of speakers ?

I don't think we can be sure. Certainly we cannot unless we experiment with several members of the family of interactions between the variants. A priori I don't know, and in this field there is no way of knowing, whether or not '% difference between F and R on common nouns' is a better discriminator amongst people for whom the measure varies than '% difference between F and R + FR on common nouns'.

For any particular definitional form of an interaction between variants, we can only apply it to a sample of speakers, in combination with other criteria, and watch what happens to the sample of speakers. Whether, as a result of such watching, we turn out to be able to decide which definitional form of a criterion is 'better', or 'more useful' may well prove to be an open question.

B (ii): Intuited extensions to criteria.

In 1972 my reaction to the inclusion and definition problems was optimistic (Pellowe et al. 1972). The number of dimensions was increased without reference to further data, by subdividing categories of variation covered by criteria discussed up until then, and by including dimensions able to represent variation in the paralinguistic systems of tension, voice qualifiers and voice qualifications (Crystal 1969; Crystal & Quirk 1964). These additions were made on the basis of native (Tyneside) and non-native intuitions in the face of findings about 'educated' speech – especially Crystal (1969) [N].

[N] Thus any co-occurrence pattern which is of high or low frequency in (some type of) 'educated' speech is a potentially good criterion – because the frequency in localised varieties may be the opposite (cf. B,C15). However this potential is complicated by the fact that though the systems (as defined by Crystal & Quirk 1964) are logically independent, there are some associations between terms in different systems which may be empirically dependent in highly similar ways in very different varieties.

Furthermore, other additions were made on the basis of guesswork: what would this sound like in the mouth of **x**, or **y**, or z (B,C18) ?

In this fashion the number of criteria was increased from 22 (Fig. B12) to 61 (Pellowe et al. 1972a). They are distributed in groups as follows (criterion numbers as in Pellowe et al. 1972a):

(i) 4 criteria dealing, in various ways, with the length of TUs (criteria 295-298 inclusive);

(ii) 28 criteria dealing with interactions between tone and form-class (criteria 299-323 and 446-448);

(iii) 12 criteria dealing with interactions between tone and pitch-range (criteria 324-335);

(iv) a miscellaneous group of 17 criteria dealing with various interactions in the systems of prominence, tempo, rhythmicality, tension, voice quality and voice qualification (336-352).

Those criteria from Pellowe et al. (1972a) which we shall have cause to consider are those in groups (ii) and (iii). Each group was derived from a matrix of totals for each speaker. The first matrix is of the cooccurrence between tones and specified word-class categories; the second of the cooccurrence between tones and terms from the simple system of pitch range.

	Ø	n	N	adv	padj	fadj	pron	tag	verb	tot
Ø		G2	G3	G4	G5	G6	G7	G8	G10	G9
F	H1	H2	H3	H4	H5	H6	H7	H8	H10	H9
R	J1	J2	J3	J4	J5	J6	J7	J8	J10	J9
FR	K1	K2	K3	K4	K5	K6	K7	K8	K10	K9
RF	L1	L2	L3	L4	L5	L6	L7	L8	L10	L9
FpR	M1	M2	M3	M4	M5	M6	M7	M8	M10	M9
RpF	N1	N2	N3	N4	N5	N6	N7	N8	N10	N9
L	P1	P2	P3	P4	P5	P6	P7	P8	P10	P9
tot	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q10	Q9

(column 9 includes column 10)

Notes on matrix:

(a) G-row refers to totals of word-class categories not carrying nuclear tone (or a part thereof).

(b) 1-column refers to totals of nuclear tones not carried by the specified word-classes.

(c) Notation of matrix. n common noun; N proper noun; adv participating in adverbial phrase (including time, place, manner adverbials); padj Fries' form-class operating as modifier (i.e. excluding Halliday's modifier classes determiner, ordinal, intensifier, nominal) – thus in "those three very pretty Newcastle girls" the only padj is pretty; fadj Fries' form class operating as complement of a verb; pron pronoun as subject or object; tag always clause final, question form repetition (in referential terms) of S & P elements of preceding sentence, with positive or negative, constant or reversed, polarities [he is, is he ?; he is, isn't he ?; he isn't, is he ?; he isn't, isn't he ?]; verb modal, auxiliary or lexical element of a finite verbal group.

Criteria derived from the above matrix were as follows. (The numbering of the criteria duplicates that of Pellowe et al. (1972a) - cf. also Appendix A, above. The expression of the ratios follows arithmetical conventions.)

299 (H9-J9/Q9-G9)100 % difference between F and R nuclei.
300 (J9/Q9-G9)100 % all nuclei which are R.
301 (P9/Q9-G9)100 % all nuclei which are L.
302 (M9/Q9-G9)100 % all nuclei which are F(pR).
303 (H9-(J9+P9)/Q9-G9)100 % difference between F and the sum of R and L.
304 |((H2/H9-H1)100)-((J2/J9-J1)100)| the absolute difference between the % of common noun Fs and the % of common noun Rs.
305 (J4+J7/J9-J1)100 % of Rs on adv. and pron.
306 (H7/H9-H1)100 % of Fs on pron.
307 (K2/K9-K1)100 % of FR on common nouns.
308 (L2/L9-L1)100 % of RF on common nouns.
309 ((K3+L3)/(K9-K1)+(L9-L1))100 % of RFs and FRs on proper nouns.
310 (K4+K7/K9-K1)100 % of FR on adv and pron.
311 (P8+J8/(P9-P1)+(J9-J1))100 % of R and L on tags.
312 (H8+K8/(H9-H1)+(K9-K1))100 % of F and RF on tags.
313 (P2/P9-P1)100 % of Ls on common nouns
314 ((P3+P4+P5)/P9-P1)100 % of Ls on N., adv. and padj.
315 (G2/Q2)100 % common nouns <u>not</u> carrying nuclear tone.
316 (Q1/Q9)100 % tones <u>not</u> carried by these grammatical categories.
317 (J4+K4/Q4-G4)100 % nuclear adv.s with R and FR.
318 (H7/Q7-G7)100 % pron. with F
319 ((|H5-M5|)/Q5-G5)100 Absolute difference between F and F(pR) on padj, % padj.
320 ((H2-J2)/Q2-G2)100 % difference between Fs and Rs on common nouns.
321 (H6/Q6-G6)100 % fadj. with F.
322 (M6/Q6-G6)100 % fadj. with F(pR).

323 ((J3+P3+J7+P7)/(Q7-G7)+(Q3-G3))100 % N. and pron. carrying R and L.
446 (G8/Q8)100 % tags without tone.
447 (H10/H9-H1)100 % nuclear verbs with F.
448 ((H10/Q10-G10)100)-((J10/Q10-G10)100) % difference between nuclear verbs with F and those with R.

	Ø	extra high booster	high booster	booster	cont-inuance	drop	low drop	tot
Ø		R2	R3	R4	R5	R6	R7	R8
F	S1	S2	S3	S4	S5	S6	S7	S8
R	T1	T2	T3	T4	T5	T6	T7	T8
FR	U1	U2	U3	U4	U5	U6	U7	U8
RF	V1	V2	V3	V4	V5	V6	V7	V8
F(pR)	W1	W2	W3	W4	W5	W6	W7	W8
R(pF)	X1	X2	X3	X4	X5	X6	X7	X8
L	Y1	Y2	Y3	Y4	Y5	Y6	Y7	Y8
tot	Z1	Z2	Z3	Z4	Z5	Z6	Z7	Z8

(The convention for coding this matrix is that when a tone co-occurs with a term from the simple pitch range system, the latter must <u>immediately</u> precede the nuclear syllable.)

324 (Z1/Z8-R8)100 % nuclei <u>not</u> accompanied by some term from the p-r system.
325 (Y8-Y1/Y8)100 % Ls co-occurring with a p-r term.
326 (Y2+Y3+Y4+Y5/Y8-Y1)100 % Ls co-occurring with any p-r term which is non-drop.
327 (T8-T1/T8)100 % R with p-r term.
328 (T6+T7/T8-T1)100 % R with p-r term which is drop/low drop.
329 (S8-S1/S8)100 % F with p-r term.

330 (S2+S3+S4+S5/S8-S1)100 % F with p-r term which is non-drop.

331 ((U8-U1)+(V8-V1)+(W8-W1)+(X8-X1)/U8+V8+W8+X8)100 % all FR, RF, FpR, RpF tones which co-occur with p-r term.

332 (U2+U3+U4+U5+W2+W3+W4+W5/(U8-U1)+(W8-W1))100 % all FR and RF with a p-r term which is a non-drop.

333 (V6+V7+X6+X7/(V8-V1)+(X8-X1))100 % all RF and RpF with a p-r term which is drop low drop.

334 (Z2-R2/Z2)100 % extra high boosters which immediately precede a nuclear tone.

335 (Z6-R6/Z6)100 % of all drops which immediately precede a nuclear tone.

I now completely repudiate my former willingness to add criteria to the system which are not directly derived from relevant data. I do so on the basis of two considerations:

(a) the stability of intonational criteria, and

(b) the optimality of intonational criteria. And these two are interdependent.

<u>(a) Stability</u>

We require intonational criteria to be <u>stable</u> in respect of new subsamples of speakers drawn from the same population. (This requirement reflects much of the preceding discussion, particularly that concerning the simultaneity of realisational variability and varietal variability, the low recognisability of intonational variants, and the weakness of our intuitions about them.)

What exactly is it which constitutes stability in an intonational criterion ? A stable intonational criterion is one which having represented or dispersed a subsample, n_i, of speakers in some particular way, will represent or disperse a further subsample, n_j, of speakers from the same population in a compatible way. 'Compatible' here will vary in its exact signification depending on the nature of the dispersion. For example, if the dispersion of n_i is linear in a straggly sort of way, then a compatible dispersion of n_j is one along a line with roughly the same equation; an incompatible dispersion of n_j is one , for instance, at right angles to that of n_i. The reason for using a word as vague as 'compatible' is that it is perfectly possible that one

<u>part</u> of n_j will be in accord with the dispersion of n_i and the <u>other</u> part will, in extending the dispersion of n_i, change it from a linear to a curvilinear dispersion. This would not be incompatible in the sense in which I am using the term, and we would want to say that such a criterion was stable on n_i, n_j, especially if it dispersed a further subsample n_k in a compatible way. And so on.

We have already come across what looks like a wildly unstable criterion, in the comparison between Pellowe's (1970d) and McNeany's (1971) empirically derived criteria (viz. criterion (10) % R with Ø p-r term (Figs. B9 & B11)). [N]

[N] I set aside, for the present discussion, the arguments about the oddity of informant 'NL' (see above p. 192ff).

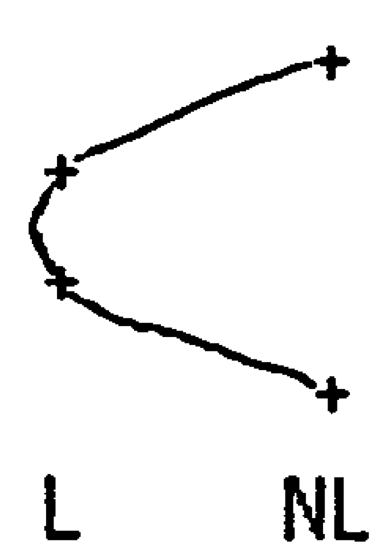
The values make the point:

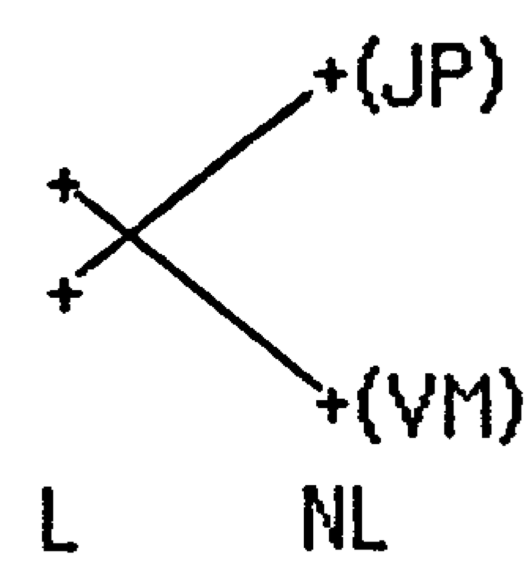
ØR, %R:	VM	JP
NL-varieties	33	74
L-varieties	60	42

Even if we do not know what it means of a criterion to say that it is stable, it is, at least, a method of shoring up our very shaky intuitions. We shall be inclined to accept criteria which are stable and to reject those that are unstable. [N] It will surely be a minimum requirement for the stability of a criterion that it be derived from relevant data.

N Two complicating points may be raised in objection:
(1) If we don't know what stability in a criterion means, then stable criteria may not be the best way of representing the variants (even if they are the best way of representing the subsamples of speakers). The answer to this is surely that if we don't know what 'stable' means, then neither do we know what 'best' means. Both forms of ignorance derive from a third: not knowing what intonational variability is like distributionally.
(2) The competing trends which are said to make ØR,%R a wildly unstable criterion are not two linear trends, but one curvilinear trend,

i.e. (a) rather than (b)





I have nothing against this objection as such - it satisfies the notion of a compatible assimilation of $\mathbf{n_j}$ to $\mathbf{n_i}$ - although the change brought about by the addition of $\mathbf{n_j}$ stretches my characterisation of 'compatible' rather drastically. But note that the argument of this objection has highly significant linguistic consequences. It means that the criterion is projecting two radically different forms of non-localised variety, with all localised varieties intermediate between them. This is certainly not impossible. But I would expect to find such (curvilinear) dispersions on terms from the systems of tension and voice quality, rather than on those from the systems of tone and pitch range, because as Crystal & Quirk (1964) suggest, the former systems are less linguistic than the latter and hence, I imagine, they will be more open to affective (or, possibly, random) variability (B,C19).

(b) Optimality

The second consideration which persuades us to obtain criteria empirically, and not by intuited or comparative extension, is that of the optimality of intonational criteria. The kind of optimality I am referring to here is methodologically imposed, rather than empirically demonstrable (cf. (1) in the previous note).

I will say of an intonational criterion that it is optimal if:

(i) its definition integrates it with distributionally related intonational features or structures, and

(ii) it maximizes the divisive effect (on the samples of varieties) of its varying feature (or primary varying feature if it has more than one).

Notice that both these characteristics are matters of degree, and that optimality is not therefore present/absent in some criterion, but rather more/less. Further, notice that both characteristics require the repeated comparison of competing definitions of each criterion. What (i) means is that if we find a higher frequency of variant **x** in L-varieties compared to NL-varieties, and a lower frequency of variant **w** in L-varieties compared to NL-varieties (especially when **x** and **w** are in the same, or closely related, intonational systems), then we should try to define a criterion incorporating **x** & **w** relationally.

(A good example is the attempt in criterion (1) Fig B 12, viz. the relative %s of F:R, F:L, R:L.)

What (ii) means is that a criterion is no good if it doesn't separate people in some way which looks plausible of interpretation (even if only in conjunction with other criteria). Thus a criterion which represents a whole sample of speakers as a diffuse clump or as a uniform dispersion in two dimensions is not discriminating varieties in interpretable ways, if at all. [N]

[N] There is a problem of cyclicity here. 'To maximize the divisive effect' is not just a matter of defining in the most appropriate way, but is also dependent on feeding the criterion with an appropriately varied sample of varieties for dispersion. This is further demonstration of the need to arrive at criteria iteratively from new sources of data.

The function of notions such as the stability and optimality of intonational criteria is to ensure that we continuously and appropriately interrogate a satisfactory data base, both in terms of speakers and in terms of criteria. The process of determining, from the network of co-occurrents of a given varying feature in a sample of speakers, the manner in which that varying feature must be defined as a criterion is complex in ways which will be depicted in a moment. [N]

[N] It would be worth developing programs for the interactive computation of these complexities and pilot work has shown that such programs are feasible. I am very grateful to Val Jones for discussion of these problems, and for her algorithmic sketches of the feasibility of such computations. (See below Appendix D.)

I now turn to an exploration of the behaviour of some of the criteria in characterising the realisations of samples of interviewed speakers. It is important to stress that our interest is focussed as much upon what the varieties (speakers) do to the criteria as upon what the criteria (dimensions) do to the varieties.

C. Application of criteria to a sample of speakers: the appearance of structure. N

N Of course the word 'appearance' here is three ways ambiguous, meaning roughly: emergence/perceptible form/erroneousness. This seems to me to be useful.

First, I try to characterise typical forms of varietal (or lectal) intonational variation by considering the frequencies and co-occurrence distributions for three classes of features in the speech of twenty informants. (What follows draws substantially on Pellowe & Jones 1977.) The informants are members of stratum 4 of a random sample of speakers from Gateshead on Tyne (Pellowe et al. 1972: 21ff.). N

N This random sample was drawn using a prestratifying factor of 'rateable value per dwelling per polling district'. Stratum 4 is the lowest but one. The sample was designed, and the interviews were conducted, by Vince McNeany.

The ages of the informants range from 17 to 70, there are 9 women and 11 men, distributed as in Figure B13.

(Letters are mnemonics for informants, bracketted numerals show socio-economic class according to: V unskilled manual, IV skilled manual and routine non-manual, III lower supervisory, II higher supervisory (Hall & Jones 1950).) All informants left school at the legal minimum for their dates of birth, except N and Mc, who had received full time tertiary education (polytechnic and college of education respectively). All informants had lived continuously on Tyneside, and so had their parents, except Ch (born in Ulster of Ulster parents, but moved to Tyneside shortly after birth) and Gr (born of rural-living parents in Tyneside's hinterland). This group of informants (characterised in Fig. B13) will be referred to collectively as Group I (Gp I).

The three classes of linguistic features here discussed are tones, selected form-classes, and terms in the system of pitch range. Tone is represented in a system of eight terms: F (fall), R (rise), FR (fall-rise), RF (rise-fall), FpR (fall plus rise), RpF (rise plus fall), L (level), and Ø (zero, i.e. no tone). Form-class is represented by one of nine terms: n (common noun), N (name), adv (adverb), padj (premodifying adjective), fadj ('adjective' as complement), pron (pronoun as subject or object), tag (clause final; includes 'phatic' (you know, I think) as well as 'syntactic' (he is, is he; he is, isn't he) types)*, vb (finite verbs), Ø (zero, i.e. some form class other than these). The simple system of pitch range is represented by a system of seven terms: ehb (extra high booster), hb (high booster), b (booster), c (continuance), d (drop), ld (low drop), Ø (zero, i.e. no marked pitch range feature). Pitch range has been described in detail by Crystal & Quirk (1964), and by Crystal (1969). Differences of pitch which occur within an utterance may be divided into those which are dynamic (i.e. which constitute glides) and those which are static (i.e. comprise contiguous syllables which are of different pitch). [N]

[N] Level tone is not dynamic in this sense, but excellent reasons are adduced for its inclusion as a tone by Crystal (1969: 215-7).

These latter, static, pitch differences are the elements of the pitch range system. Elements are recognised according to the nature and degree of the relationship between the pitch of the syllable under consideration and that of the previous pitch-prominent syllable.

The co-occurrence of any feature, of these three sets of features, with any others, can thus be represented by an ordered three element code. There are 504 possible co-occcurrence codes for these three systems, derivable from the following matrix:

Tone	(code)	Form class	(code)	Pitch range	(code)
Ø	G	Ø	G	Ø	G
F	H	n	Q	ehb	A
R	J	N	R	hb	B
FR	K	adv	S	b	C
RF	L	padj	T	c	D
FpR	M	fadj	U	d	E
RpF	N	pron	V	ld	F
L	P	tag	W		
		vb	X		

Thus for instance, a RF occurring on an exclamatory word (e.g.well) would be coded LGG; a name preceded by a booster, GRC; a fall on an intensifier preceded by a continuance, HGD, and so on. Note that occurrences from any of the systems which do **not** co-occur with any terms from the other two systems are coded nonetheless (e.g. GXG, GGE, KGG, etc. But GGG is not coded.). Clearly, adequate information on occurrence is a necessary prerequisite for any satisfactory account of co-occurrence.

The data comprise 4066 TUs (mean of 204 TUs per informant; highest number of TUs for one informant 273, lowest 142). Occasionally comparison is made with the results of Quirk et al. (1964). Their sample (of 10 speakers in two panel discussion groups) yielded 1880 TUs. The informants discussed here however were speaking in a loosely-structured interview (by the standards of e.g. Oppenheim (1966)). Results of Crystal and Davy (1969) and of Crystal (1975:96 ff.) show that the realization patterns of prosodic and paralinguistic features are likely to change under such changes of interactive purpose. In spite of this, the comparisons prove useful. [N]

[N] I am indebted to Vice McNeany for his analysis of these informants.

Consider first the gross frequency distribution of tones in Figure B14. (The criteria depicted correspond to Nos. 300 to 302 in Pellowe et al. (1972a) (cf. above, p.211), and criterion (2) in Fig. B12.) Against the trace for Quirk's (1964) sample are given those for two Tyneside samples (T_1 and T_2), each being a random half of Group I – the twenty speakers here considered. (T_1= He, Ga, Wi, S, Wa, N, O, Cl, T, Cr; T_2= the rest.) The match between the two random halves is good, certainly in terms of the overall differences from the trace for Quirk et al.'s (1964) sample (henceforth abbreviated to Q). There is a sizeable difference between non-localised varieties (Q) and Tyneside localised varieties (T_1, T_2) in respect of the relative fractions of Fs, RFs, and Ls. Such differences as these may be thought of as diagnostic of these two classses of variety.[N]

[N] It will be recalled from my discussion of the model (see above, Ch. 3) that dimensions can be thought of as dividing the population up into groups. Clearly different dimensions will produce different numbers of groups of differing sizes. In other words some dimensions diagnose more abstract groupings and some diagnose less abstract groupings in the population. More formal treatments of the notion of the diagnostic function of linguistic features appear in Pellowe et al. (1972) and Pellowe (1973), and in general mathematical terms in Sneath & Sokal (1973: 381-408) and Kendall & Stuart (1968: 314-336). For a general summary of the problems, techniques and goals of taximetrics and taximetric diagnosis see below, Appendix D.

This is certainly true at this level of representation, i.e. gross percentages for whole samples. But here we must take into account the important work of Garvey & Dickstein (1972) who demonstrate that radical changes of relationship between dependent and independent variables can be brought about by changes in their form of representation.

We can choose a different level of representation for the percentage distribution of tones by, for instance, choosing individuals rather than whole samples. When we do this - (Figure B15) - we find that the component members of a distribution like T_1 show considerable differences of tone distribution, but that, even with such an unmatched group of ten speakers, these differences are patterned. Such patterning of individual differences of tonic distribution within T_1, given the fit between T_1 and T_2, suggests that the differences may be stable. This possibility is examined in subsequent pictures.

Consider first the relative fraction of Fs and Rs represented for the sample in Fig. B14. Non-localised varieties (Q) appear to have a much larger percentage difference between Fs and Rs (25%) than do localised (T_1, T_2) varieties (10%). If we specify this (i.e. % difference between Fs and Rs) as a dimension we might expect to find a distribution of speakers in which those with high values were less localised than those with low values (B,C20).

Figure B16 shows the fate of all speakers (T_1, T_2) plotted by age on the difference between the number of Fs % tones and the number of Rs % tones.

Four elements of this picture are worth emphasising here.

First, the dimension is clearly sex-differentiating: men have high values on the dimension, women have low ones. The expectation expressed above - that the dimension should discriminate localised from non-localised varieties - is not satisfied by this picture. All the speakers in Q were men amongst whom the mean difference between Fs and Rs was 25%. [N]

N Notice that the men in the Tyneside sample have a similar mean % difference. This similarity is probably spurious for two reasons: (a) there is reason to imagine that panel discussion (Q), and informal interviews (T_1, T_2) would elicit different types of realisation pattern (Crystal & Davy 1969), in which case distributional sameness may very well indicate difference of underlying system;
(b) comparison of the university educated speakers in Q with the most educated speakers in T (Mc, N) shows that the latter have a far higher % diff. F/R than the mean Q value and are therefore likely to be producing realisations of a different system.

Rather than being diagnostic of the localised/non-localised nature of the variety, the dimension seems to be primarily diagnostic of gender.

Secondly, however, the four men having the highest values (Mc, N, Ch, Gr: >28% - see enclosed area on Fig. B16) are the ones differing from the rest of the T sample in either education (Mc, N), or nuclear family residence pattern (Ch, Gr).

Thirdly, for the remaining men, there is a strong tendency for their realisations to comprise about 20% more Fs than Rs, irrespective of age.

Fourthly, amongst women there is an age trend which indicates that younger women are realising Rs in more and more TUs in which their elders would have realised Fs. (See the dotted area on Figure B16.) This is a trend which seems to be socially significant for members of the speech community. [N]

[N] It seems to be a behaviour emulated by Ar, who in terms of her age should have a value of +15% or so, but in fact has a value of -26%. We know independently that Ar goes dancing, listens to pop-music, and reads teenage magazines. But given the complex nature of the dimension (% diff. F/R), what might be involved in the mental processes underlying such social salience (if such it be) is somewhat hair-raising. Cf. in this respect Bickerton's suitable satire (1971:460-1) of Labov's notion that speakers must somehow keep a cumulative total of the variants they use. Notice that in the model I am advancing no such direct psychological reality is proposed. See above Chs. 2 & 3 for discussion of the necessary, but indeterminate, distinction between $_{h}$Vs and $_{a}$Vs.

In the light of the remarks above about interaction between criteria, it would be peculiar if the trends I have just discussed were without their connections in other parts of the tonic system (B,C21). Figure B17 pictures the detailed relationships, for individuals, between the frequencies of Rs, Fs, and Ls (the three most frequently occurring tones) for the T sample as a whole.

There are certain clear patterns of frequency distribution here which are represented qualitatively in Figure B18.

For women speakers (see Fig. B18) there are here two types of subvariety of tone-distribution:

Type I, in which the % of Rs is considerably greater than the % of Fs [4 out of 9 speakers] (cf. Fig. B16 and the discussion of it above, p. 228ff.);

Type II, in which the reverse is the case - the % of Fs is considerably higher than the % of Rs [5/9 speakers].

Each of these types has two subtypes:

(a) in which the % of Rs and Ls is about equal,

(b) in which the % of Fs and Ls is about equal.

Amongst men variation of frequency pattern of this kind is much less widespread - there is considerable uniformity of the frequency % of R and F tones. There is one tone frequency pattern in common between men (a) and women (IIa), which, as a pattern, represents 13/20 members of this sample. It is worth noting that the three women having this common pattern are in socio-economic class V, which lends support to the argument advanced in the penultimate note. This shared pattern however, is quantitatively different for men and for women (see Figure B19). First, the mean % diff. between Rs and Fs is lower for women (13%) than for men (25%) – in spite of the results of Fig. B16. Secondly, the % diff. for women falls within the range of the % diff. for men. If we were to choose to use such tone frequency patterns as a dimension – each pattern possibly being represented as some part of one dimension – it would not be a dimension which discriminated men from women according to any simple principle. Where there is a qualitative difference of pattern (F, I(a) against M(a), for example) the discrimination will be all-or-none. But when the patterns are qualitatively the same (F, II(a) against M(a)) the discrimination can only be probabilistic, and the probabilities inhering in the pattern would have to be continuously monitored from the samples of speakers being considered.

The kinds of dimensions we have been considering as possibly suitable ones for an optimal variety space (Pellowe et al. 1972:10;

and see above Chs. 2 & 3) may be distinguished as to whether they are simple or complex. A complex dimension is one that directly incorporates a relationship between independent features (which may also be independently varying for different varieties in the speech community).[N]

[N] Note that complex dimensions satisfy the first part of the definition of optimal intonational criteria (cf. above, Appx B, section B(ii), p. 218).

I have already discussed the behaviour of one such complex dimension - the % diff. between Fs and Rs - in respect of age (cf. Fig. B16). Simple dimensions reflect single varying features.

In the next figures I examine the interaction between tonic frequencies on simple and complex dimensions. I concentrate mainly on Fs, Rs, and Ls, since these constitute 67% of TUs in the whole sample, ranging in individuals from 83% to 55% of their totals. (Cf. Q in which the same tones account for 75% of the total TUs.)

In Figures B20a, B20b and B20c, I return to a treatment of Fs and

Rs by themselves. The horizontal axis is the same in both figures and is the same complex dimension we have already seen in Fig. B16, notably the % diff. between Fs and Rs. In Fig. B20a, the % of Fs is plotted against this dimension; in Fig. B20b the % of Rs is plotted against this dimension. Both figures plot individuals (women encircled, men not); and, for comparison, the position of the Q corpus as a whole is given. [N]

[N] Whether the curves in these, and subsequent, figures are interpreted as regression lines or as lines representing correlation is not really material at the present point in the work (cf. Kendall & Stuart 1967: 278-9). However what was sought was the dependence of **x** (i.e. % diff. F/R) on **y** (e.g. % F) - and so one can think of the graphs as representing lines of regression. The curves were fitted by eye. Some have been checked by minimising the sum of squares (Kendall & Stuart 1967:75ff., 286-7, 346ff.) and the degrees of fit were good. In fact none of the discussion turns on the accuracy of the curve fitting, and given the nature of the sample and the variates it would probably have been misleading if it had. (Neither the curve fitting nor minimising ΣD^2 took account of Q.)

Several points in Figs. B20a and B20b need comment.

First, we can dispose of two negative points: (a) the sample is ranked in the same way on **x** in the two plots because **x** is the same in the two; (b) the fact that the regression cuts the y axis at the same value of y (21%) in both plots is a coincidence which doesn't seem to have any significance - although it is not an artefact of the level or of the kind of representation.

Secondly, we may remark the continuity of distribution of the sample on the dimensions. For increasing positive values of **x**, the corresponding increase (Fig. B20a) or decrease (Fig. B20b) in y shows relatively small deviation from linearity. This indicates that the interdependence between Fs and Rs for speakers in this part of the graphs is fairly stable.

Thirdly, though the rank order of the sample on y (% Fs) in Fig. B20a is by no means preserved on y (% Rs) in Fig. B20b, it is far better preserved than one might expect. (The rank order is most disturbed amongst speakers having 20% to 40% Fs and 5% to 20% Rs, and having about 20% more Fs than Rs; but even amongst these speakers, certain groupings are preserved.) [N]

[N] A given rank-number will include speakers who have been upwardly displaced <u>and</u> those who have been downwardly displaced. Thus, a low-to-high list of cases on some axis y_a is matched by a list, similarly constructed, of cases as they are dispersed on a comparable axis y_b :

	y_a	y_b
..	..	..
11	L	N
12	M	M
13	N	L
14	P	Q
15	R	S
16	S	R
17	Q	P
..	..	.. ;

then we have displacements of L (+2), M (0), N (-2), P (+3), Q (-3), R (+1), S (-1);
and then, in terms of the number of ranks any case is displaced by, and the number of cases so displaced, we have:

N. of ranks	No. of cases
0	1 (M)
1	2 (R, S)
2	2 (L, N)
3	2 (P, Q)

In the case of the y-axes of Figs. B20a & B20b the rank displacement totals (+ & -) for the sample are:

No. ranks speaker displaced by:	0	1	2	3	4	5	6	7	8	9	
No. speakers so displaced:	3	3	4	2	2	1	1	1	2	1	(20)

Fourthly, however, the two distributions are not the inverses one might expect given the similarities so far discussed. The asymmetries are directly relevant to the process of establishing prosodic systems of different varieties. There are two classes of facts.

(1) The slopes (rates of regression if you will) of the straight line parts of Figs. B20a and B20b are different in value. (Their difference of sign is a consequence of the form of representation.) The slope for % Fs (Fig. B20a) is 1 in 1.6, that for % Rs (Fig. B20b) is 1 in -2.5. (I.e. 1 up for 1.6 to the right; 1 up for 2.5 to the left.) Put another way, for Tyneside varieties of the kind here analysed, a unit increase in the use of Fs by a speaker will increase the ascendancy of Fs over Rs by 1.6, whereas a unit increase in the use of Rs will decrease the

ascendancy of Fs over Rs by 2.5 - though the precise values are not relied upon here.

(2) The curvilinear parts of the traces arising for negative values of % diff. F/R are different in the two figures. Consider the negative part of Fig. B20b. When speakers realise more than about 25% of their tones as rises it seems that the rate of decrease of x (% diff. F/R) drops quickly from 1 in -2.5 to 2 in -1. Put crudely, such speakers can be thought of using their increased number of rises for something other than the 'functions' of fall. [N]

[N] Statistically minded readers might wish to argue for a completely curvilinear representation of both these figures - though the data do not seem to me to justify it. The argument, however, would not be affected since it turns on the overall characteristics of the traces, not their exact equations.

The preceding supposition is strengthened by considering the negative part of Fig. B20a. Here we find that speakers who have the largest negative values on the difference between falls and rises (i.e. many more rises) do not have the lowest number of falls. [If we glance forward to the locations of D and Ar in Fig. B21 (which represents the % diff. L/R against the % L), we find further confirmation of this claim.] Rise is taking over some of the distribution of level tone in the mouths of speakers who realise more than 15% more rises than falls. It seems quite likely that such rises are functioning as low contrast alternants for localised level tone. But it is impossible to be more definite about this on the basis of this material.

Finally, the relationship between D, Ar, and Q on the y (vertical) axis of Figs. B20b and B21, and on the x axis of Figs. B21 and B22, indicates something of the distinctiveness of the tonic distributions of those women.

An account of the varieties in these data must, in terms of this complex dimension (% diff. F/R), certainly group D & Ar together and separate them from Wa, and probably from S. The rest of the sample is also grouped in various ways. There seems to be a trend for socio-economic class to rise with rising values of the complex dimension. Men have a tendency for higher positive values than women.

Figs. B20a and B20b have been discussed at some length in order to show two things:

(a) that there is structure in the variation of such simple surface phenomena as the proportions of different tones realised by a small sample of speakers;

(b) that the dependencies amongst such phenomena, when they are depicted quantitatively, can be shown to be quite complex.

In general I do not see any reason to imagine that all, or even most, of the structured variation in such features as we are discussing will correlate with social macro-variables (such as socio-economic class) (B,C22). The primary goal must surely be to find, not dimensions which correlate with such things, but dimensions which are stable for changed samples and which are capable, in combination, of eliciting the structure of the variation in those samples. Only when we know enough about such stability and such structure will we know what level of thing to look for as correlate or, indeed, whether to look for correlates at all.

In the discussion of Figs. B20a and B20b it was remarked that the curved parts of the distribution had to do with a relationship between falls and rises, and levels and rises which was varying in different speech varieties in the sample. To express this relationship further I have plotted in Figures B21, B22, and B23 the percentage difference between levels and rises against, respectively, the percentage of levels, the percentage of rises, and the percentage of falls. (They are printed together as Figure B23a for ease of comparison.)

Some of the remarks which applied in general terms to Figs. B20a and B20b apply here also. The linear distribution of the sample in Figs. B21 and B22 is similar to that in Figs. B20a and B20b. Figure B23 however, represents the sample on two competing trends. First, there is the trend which shows that there is a class of speakers for whom an increased percentage of falling tones is related to an increasing gain of level over rising tones (e.g. Ar, T, Ha, F, Gr, indicated by a continuous line). Secondly, there is the trend which shows that there is a class of speakers for whom an increased percentage of falling tones does not covary with changes in the relative fractions of levels and rises (Wa, We, Cl, Mc, N - indicated by a dotted line and supported by the location of Q). [This second trend is shown in 'pure' form in Fig. B24 below.]

The distributive effect of the axes on the sample is similar in Figs. B21 and B22, but Fig. B23 is distinct from them in this respect. At the same time, a comparison of individual speakers in the three Figures shows that the dimensions in Figures B22 and B23 are much more similar in their ranking of speakers than either set is to that in Fig. B21. (See next note.) In other words the dimensions of Figs. B21 and B22, taken singly, are much more likely to be stable as dimensions (i.e. stable ways of dividing up a sample of speakers) than those of Fig. B23. On the other hand, Figs. B22 and B23, considered together, represent interestingly systematic behaviour on the part of the variables (rather than the speakers). That is, the inverse relation between the frequency of rises and of falls, plotted against the % difference between levels and rises, is quite strongly rank preserving. [N]

[N] Rank displacements of informants on y axes of Figs. B21, B22, & B23 are as follows (y_{21}=% levels; y_{22}=% rises; y_{23}=% falls; 20 ranks):

No. of ranks speaker displaced by (pos. & neg):	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
No. speakers so displaced comparing:																			
y_{21}/y_{22}	1	1	5	1	2	2	-	1	2	1	-	1	-	1	-	-	2	-	-
y_{21}/y_{23}	2	1	5	-	-	-	1	1	2	1	1	3	-	1	-	1	-	-	1
y_{22}/y_{23}	3	2	6	2	1	-	2	2	1	-	1	-	-	-	-	-	-	-	-

Comparison of Figures B21 and B22 with Figures B20a and B20b shows further asymmetries. Unlike the situation in Figs. B20a and B20b – where negative values for the % diff. F/R showed competing curves on both % Fs and on % Rs – the complex dimension in Figs. B21 and B22 distributes the sample less complexly. The difference between Ls and Rs decreases at the rate of 1.4 for unit increase in Rs, for <u>all</u> numbers of Rs (Fig. B22). On the other hand, the difference between Ls and Rs increases at the rate of 1.5 for unit increase in levels, from about 8% L upwards.

A very complex perturbation of the tonic system is caused by the formal-functional distinctiveness of level tone in Tyneside varieties compared to level tone in non-localised varieties. One of the reasons for the complexity of this disturbance is that we do not seem here to be dealing with a system of discrete varieties. Varieties can range continuously, as can be seen from the distributions so far discussed, from one pole of a dimension to the other (B,C23). There is a range of localised Tyneside varieties, in other words, in which the distribution of levels is differentially 'ambiguous'. In some varieties the distribution seems to be equivalent to non-localised levels. In some varieties part of the distribution seems to be equivalent to non-localised levels, part to localised rises. In some varieties part of the distribution seems to be equivalent to non-localised levels, part to localised falls. In some varieties parts of the distribution seem to be equivalent to all three. Even in varieties in which the fraction of level tones is small, there is the likelihood that rises will take on a distributional omniverousness of the same kind as this which I have outlined for levels.

Such broad generalisations can be arrived at by detailed comparative scrutiny of Figs. B20a, B20b, B21, B22, B23, B23a and B24. The following heuristics retrieve a good deal of the relevant information:

(a) Consider the negative parts of the traces of Figs. B20a and B20b (in Fig. B20c for convenience) on the one hand, and Figs. B21 and B22 (in Fig. B23a for convenience) on the other, in respect of the location of the Q sample;

(b) Consider the differential ranking of speakers (on both axes) in Figs. B20a & B20b, (=Fig. B20c) and B24 on the one hand, and Figs. B21 & B22 & B23 (=Fig B23a) on the other.

As we would expect from the preceding summary (B, C24), the pictures in which explicit dimensional power is given to Ls - whether as a simple or as a complex dimension - represent varieties which may be the most localised ones on a continuum which is more or less perpendicular to the distribution of Ls in respect of non- and less- localised varieties (Q, N, Wi, against D, Ar, S, He, Ha, An, etc.). This is also clearly shown in summary fashion by Figure B25, where I plot the interaction between the two complex dimensions so far discussed, that is, the percentage difference between F and R, and the percentage difference between L and R.

It will be recalled that Figure B24 indicated that high % falls was associated with [+male] and with [+higher socio-economic class] - reminiscent of Figs. B20a & B20b (=Fig. B20c). Figure B25, in addition to confirming previous remarks about the incursion of rises on the distribution of localised levels in the mouths of younger female speakers, indicates two different socio-economic trends.

For socio-economic group V there is a tendency for increasing values of x (% diff. L/R) to be associated with increasing values of y (% diff. F/R). Crudely, class V speakers having a certain proportion of levels to rises will have a similar proportion of falls to rises. For socio-economic group IV however, there is a tendency for increasing values of **x** (% diff. L/R) to be associated with constant values of **y** (% diff. F/R). In other words, whatever the proportion of levels to rises in group IV varieties, the frequency of falls over rises is more or less constant. [N]

[N] But given that the setting of the boundaries of the socio-economic classes (IV & V) must be entirely arbitrary from the point of view of tonic distributions (and may even contain a smidgin of arbitrariness from the point of view of the classes themselves), it would be wise not to rest too much belief in such a regularity.

In terms of the dimensions alone, that is excluding considerations of gender, socio-economic class, age and so on, Figure B25 projects three types of intonational variety from these twenty speakers (Gp. I):

(a) those having more rises than either falls or levels (D, Ar, S);

(b) those having about equal fractions of rises, falls and levels (T, He, Wa);

(c) those having more falls than rises and more levels than rises, in roughly equal degrees (BW, Ha, An, O, Cr, F).

Notice that the location for the whole Tyneside sample is in this third category. [N]

[N] There are outliers, of course. It might not be entirely foolish to expect that the discontinuities represented here would be removed by the use of a larger sample. (See below, section D.)
Notice however, that the notion of 'variety' (whether $_{h}V$ or $_{a}V$) is not discomposed by the likelihood of our finding continuities along parts or wholes of a dimension (Pellowe et al. 1972; Pellowe 1973, & see above Chs. 2 & 3). What distinguishes one group of varieties from another is the differentials amongst the **sums** of the different sub-parts of dimensional continuities; **not** the necessity of discreteness from all other groups of varieties on some stated number of dimensions.

When, instead of scattering the sample on two partially dependent complex dimensions as in Fig. B25, we scatter the sample on two independent complex dimensions, as in Figure B26, there are some interesting consequences. Figure B26 is a plot of the % difference between levels and rise-falls (**x**), against the % difference between falls and rises (**y**). [Fig. B14 shows that RFs have the fourth most significant difference of tonic distibution in localised varieties, compared to Q.]

First, these dimensions do not disperse the speakers with anything like the same amount of interdependence as we find in

Fig. B25. That is, there is no regression in the sample such that we are encouraged to say that such-and-such a range of values of x is predisposed to be associated with such-and-such a range of values of y. Broadly speaking, for increasing ascendancy of falls over rises, the ascendancy of levels over rise-falls is not predicted.

Secondly however, if we compare the groupings in Fig. B25 with those in Fig. B26, there are several preservations:

D, Ar, in Fig. B25 remain together in B26;

T, He, Wa, in Fig. B25 lose He to BW & Ga in Fig. B26;

An, Ha, F in Fig. B25 lose F in Fig. B26;

We, Cl in Fig. B26 have lost Mc from Fig. B25, but not by much.

Thirdly, when we compare the ranking of the speakers on the x axes of Figs. B25 (% diff. L/R) and B26 (% diff. L/RF), it shows considerable stability. N

N Rank preservation of sample by **x** axes of Figs. B25 & B26 – the **y** axes being the same in both pictures rank the speakers the same way – :

No. ranks speaker displaced by (+ & -):	0	1	2	3	4	5	6	7	8	9	10
No. speakers so displaced:	2	5	3	2	1	-	2	-	2	2	2

(21, includes Q)

The speakers with the largest displacements of rank ($x_{25}:x_{26}$) were D (+9), Ar (+10), S (+8), N (-6), Ch (-10), Q (-8), We (-6). It is notable that five of these seven have very high positive or negative values on the % diff. between falls and rises. This indicates that comparisons of other parts of the system of tone (between such people and those with less extreme values) are likely to be unstable for the type of two-dimensional pictures we have to deal with here.

Finally, although the group at zero in Fig. B25 has no numerical significance (the differnce between 0 and 1 being of no more interest arithmetically than the difference between 18 and 19) it <u>does</u> have linguistic significance. That is, a tonic subvariety having roughly equal numbers of rises, falls, levels (Fig. B25) and rise-falls (Fig. B26) is one of some interest. I return to these matters below.

What are the interactions between nuclear selections and nuclear locations in terms of form class ? Figure B27 shows the gross percentage distribution of all nuclei on form classes, and compares the distributions of the two Tyneside subsamples (T_1, T_2) with that of Q.

Unlike the distributions by tone alone (cf. Fig. B14) it appears that, apart from the differential distribution on common nouns, the form-class distribution will not provide us with a dimension for discriminating between localised and non-localised varieties. Dissection of the common noun part of this distribution by age, gender and socio-economic class for the Tyneside sample showed no trends of any interest. The difference between the mean values of T and Q is significant, but is it not a significance correlated in any way with these three macro-variables. It is a difference which probably has more to do with the distinction between panel discussion and informal interview than with lectal subvarieties. In other words we are probably here dealing with a case of realisational rather than varietal variability. Independent evidence for this view, though slight in quantity, comes from the use of Tyneside localised varieties in spontaneous monologue and informal conversation (i.e. the data of Pellowe 1970d). That data yielded a co-occurrence frequency between common nouns and all nuclei of 44%. (The value for Q was 45%.) The general supposition is confirmed by plotting the percentage of nuclei on common nouns (**x**) against the percentage of common nouns <u>not</u> carrying tone (**y**). Such a plot shows the whole sample to be clumped together quite closely with a mean value on **x** of 23% and on **y** of 43%. That is, the dependency between the number of non-tonic common nouns and the fraction of tones carried by common nouns does not significantly vary in the sample here considered. Given the complex dependencies we have seen amongst elements of the tonic system, this can only mean that the common noun differences given in Fig. B27 are not varietal.

Comparison of the results of Quirk et al. (1964) and of Pellowe (1970d) shows that of all the form classes taken account of by them - considering the distributive effect of form class on tone - adverbs are the most likely to distinguish localised from non-localised varieties. Pellowe (1970d) found that a significantly higher % of Rs and FRs were realised on adverbs in localised varieties than in non-localised varieties. Figure B28 pictures the sample distibuted by age and sex on the % nuclear adverbs realised as Rs and FRs. No interesting regularities are represented for age in the sample as a whole. For women however, though it is an unexpected and complex finding, and on this quantity of evidence certainly cannot be generalised from with any reliability, there is a very interesting distribution. The values of the dimension are parabolically distributed. The two curves represent: (a) women from all available age groups with the highest values, (b) women from all available age groups with the lowest values. The change of direction occurs at around the age of 40. In other words there seems to be a differentiation - in terms of the % nuclear adv.s having R and FR - between on the one hand the subvarieties of young and old speakers (high), and on the other hand the subvarieties of middle-aged speakers (low). Clearly a larger number of female speakers is needed for each age-group before we can have any certainty about this finding (see below, section D), and even if further data confirmed the finding, it would remain a difficult one to interpret.

Plotting % nuclear adverbs carrying Rs and FRs against the % nuclear adv.s carrying Fs (not depicted) shows a regression similar to that of Fig. B24. Men tend to have a high frequency of nuclear adverbs which are realised as falls, and a low frequency which are realised as Rs and FRs. Women show the converse tendency. Equally interesting however, is the indication that amongst men there is an age trend. Younger men tend to have a higher fraction of nuclear adverbs with Fs and a lower fraction with Rs and FRs. Older men tend to have the converse pattern.

So far we have only considered one form class - adverb - in any detail, but it is clear that tone and form class interact with complex dependencies just as was the case of elements within the tone system. In tabular form, Figure B29 gives the % distribution of the three commonest tones (F, R, L) on common nouns, names and adverbs. Figure B30 pictures the sample on the first three rows of Fig. B29 representing age and gender. Figure B31 (column 1) lists the qualitative patterns of Fig. B30 by informant. It also, in columns 2 & 3, lists patterns from figures like Fig. B30 (but not reproduced here) which represent the distributions of Rs and of Ls on n.s, N.s, and adv.s. In Figure B32a I have taken the patterns of Fig. B31 and listed together those informants who share the same patterns of tonic distribution across these three form classes. And in Figure B32b I have represented these similarity lists in a two-dimensional grouping in so far as this is possible.

Five points are worthy of note.

First, notice that the patterns of distribution of Fs on n.s, N.s, and adv.s are much less variable in this sample (12, 7, 1 for the three pattern types (cf. Figs. B30, B31, B32a)) than they are for Rs (8, 6, 6 for the three pattern types (cf. Figs. B31, B32a)).

Secondly, however, the available patterns of distribution on n.s, N.s, and adv.s are apparently the same for both Rs and Fs. Varieties differ as to whether, for falling tones for example,

either (a) the number on n.s and on adv.s is about the same and greater than the number on N.s;

or (b) the number on n.s and on adv.s is about the same and smaller than the number on N.s;

or (c) the number on n.s is greater than the number on adv.s which is greater than the number on N.s.

Varieties (though not speakers) differ in exactly parallel ways in respect of rising tones.

Thirdly, though the relative commonness of the patterns of distribution of Rs (8, 6, 6) is similar to that for those of Ls (9, 5, 6), the actual patterns of distribution of Ls are different. Fourthly though, as we would expect, the groupings of speakers given by Fig. B32b differ from those in Figs. B25 and B26, there are interesting preservations. Two outliers of Figs. B25 and B26, N and Wi, are now loosely associated, and their secondary association here with Cl & Ch (of the Cl, Ch, Ga, T group) is supported by Figs. B25 and B26. Figure B32b shows a close association between BW, Cr, & Ha and this is paralleled in Figs. B25 and B26.[N]

[N] It will be clear by now from various remarks both about the nature of complex dependencies between the kinds of dimensions here being discussed, and about how the variety space must be thought of as working, that different dimensions are not expected to group people in the same way.
Our **disproportionate and inappropriate** pleasure in seeing that they occasionally do group people in the same way is an indication of how our capacity to think about pattern falls far short of our ability to operate with (behave in terms of) extensively complex pattern (e.g.'family resemblance' - cf. Appendix D). That is tantamount to saying that left-hemispheric

conscious analysis [sc. for right-handers] can **never** match right-hemispheric non-conscious 'analysis', or pattern 'apprehension', or the creation of 'gestalts' (B, C25).

Finally, in this partial sketch of the structure of the intonational variability of twenty speakers, I consider some interactions between the systems of tone and pitch range. Fig. B11 indicated that the most widely varying interaction between these two systems (in terms of the data there represented) is booster+level (BL). (The values are high for localised Tyneside varieties, low for non-localised varieties.) Figure B33 gives two accounts of the relationship between Ls and the simple system of pitch range. The first trace (a-a) is linear by definition. It takes the number of level tones co-occurring with a pitch range feature and plots speakers in terms of the number of Ls co-occurring with a 'generalised drop' against the number co-occurring with a 'generalised booster'. Low drops and drops all count, for this picture, as generalised drops. Extra-high boosters, high boosters, boosters and continuances all count, for this picture, as generalised boosters.

For Ls co-occurring with a pitch range feature, the sample is quite strongly grouped around a value of 24% drop+level (76% booster+level). There are no marked interdependencies with social macro-variables, although values above 45% of drop+level appear to be restricted to male varieties - more cases are needed for this to be anything other than tentative.

The second trace in Fig. B33 (i.e. b-b) plots speakers on generalised boosters and generalised drops with Ls (i.e. BL & DL, as before, but as a percentage of all Ls (i.e. Ls with and without pitch range features)). High values of drop+L (but now above 35%) continue to be interpretable as male. In addition, men and women of higher socio-economic class tend to have fewer BLs - regardless of the proportion of DLs - than men and women of lower socio-economic class. Finally, there is much less variation amongst women than amongst men in respect of these co-occurring features.

On the basis of Fig. B11 it was claimed (see above, section

B(i), p. 186 ff.) that BF and BR were strongly gradient (and in the same direction) as between localised and non-localised varieties, but that BL was non-gradient (high for L-varieties, low for NL-varieties). These claims are now examined in respect of the twenty individuals in Figure B34. Here the association of generalised boosters (i.e. ehb, hb, b, c) with falls, levels and rises are plotted on the ranked points (from low to high) for BF. Four significant contrasts with Fig. B11 emerge:

(1) There is still gradience represented in BF, but since there are no NL-variety speakers in this sample of 20 - certainly An is not a NL-variety speaker - the claim made above that lower values of BF are associated with NL-varieties must be rejected. (Or reinterpreted. See below, section D.);

(2) There is no parallelism here between the gradience for BF and that for BR (contra Fig. B11);

(3) The rates of occurrence for BL and BF in Fig. B11

(i.e.	BL	BF
L-vars.	89	60
NL-vars.	33	46)

are reversed for the L-variety speakers in Fig. B34 - roughly, BF 85%, BL 50%. But the number of TUs underlying Fig. B11 is small, and we already suspect that the difference between interviews and informal conversations is significant for these variables;

(4) There is a close parallelism between BL and BR in Fig. B34 - which has nothing to do with the form of representation. At any value of BL, there is an overall mean difference between BL and BR of 11%. (This mean difference begins to rise sharply for values of BF below 83%.)

For 15 out of these 20 speakers, BL is the dominant partner of the parallelism. (Though this is true of all 9 women, I'm not very confident that a regularity of this kind has anything to do with gender.) I have no plausible guess as to why people should use 11% more levels with boosters (% levels) than rises with boosters (% rises), regardless of the relative proportions of levels, rises and falls used by the various speakers, and I defer judgement on the matter until we have more data (B, C26).

In spite of the parallelism between BL and BR, it is quite clear from Fig. B34 that neither the rates of occurrence of BL and of BR, nor their ordination, are predictable from BF. Nor are those rates and that ordination associated with gender, age or socio-economic class. Similarly, the rate of occurrence of BF itself is not associated with these macro-variables.

The negativity of these findings, however, hides a very important principle, it conceals the highly significant basis of much inference. (In one way negative findings of this sort are not negative at all.) It is important to be able to establish that features which co-occur have the capacity for different, unassociated distributions in a sample of speakers, since if they did not have this capacity, we would be less sure about the significance of associated distributions between those features when they did show themselves (B, C27).

The structure of variation: summary.

(1) There are considerable differences between the relative fractions of F, RF, & L tones in the non-localised (Q) and localised (Gp I) samples (Fig. B14); and consideration of these differences at the individual level for a random half of Gp I (=T_1) indicates that such differences might form the basis for stable criteria (Fig. B15).

(2) But plotting % difference of Fs and Rs against age for Group I does not in fact discriminate localised (Gp.I) from non-localised (Q) varieties (Fig. B. 16). Instead it shows a gender differentiation in which women have low and men high values; and amongst those women there is an age trend - the younger the woman, the greater the incursion of Rs into Fs. These facts are explored further (Figs. B17, B18, B19), but inconclusively.

(3) Comparative examination of %Fs and %Rs against %diff. between Fs and Rs shows that the dependency between the distributions of Fs and Rs is **not** symmetrical (Figs. B20a, B20b, B20c). But a person with more than -10% on the complex dimension (i.e. 10% more Rs than Fs) may not necessarily be realising 'fall functions' by R, but may be realising localised 'level functions' this way. Men tend to have higher positive values of the complex dimension than women, and s-e class may be higher for speakers having higher values (cf. also Fig. B24).

(4) There is a subset of $_{L}T$ varieties in which an increased % of Fs is related to an increasing gain of Ls over Rs; **but** in addition there is a subset of $_{L}T$ varieties in which an increasing % of Fs does **not** covary with with change in the relative fractions of Ls and Rs (Fig. B23).

Comparison of the rank-preservation of speakers on different pairs of different dimensions (e.g. Figs. B20a/B20b; Figs. B21/B22: Figs. B22/B23; Figs B21/B23) shows that for these varieties and for these dimensions, the relationships between the dimensions are stable.

The nature of the distribution of Ls in the Figs. just cited (including also Fig. B24) indicates that **Level** tone is differentially ambiguous in different $_{L}$T varieties; that is, it appears that different replacement norms for Ls (sensu Quirk & Crystal (1966)) may be attributed to different subsets of T varieties as represented by these speakers (Gp I).

(5) Pictures in which explicit dimensional power is given to Ls represent varieties which may be the most localised on a continuum which is more or less perpendicular to the distribution of Ls in respect of non- and less- localised varieties.

(6) Plotting two partially overlapping dimensions against each other (% diff. F, R and % diff. L, R) represents Gp. I on two competing trends in respect of s-e class, but probably not very

reliably. In terms of the dimensions alone, Fig B25 projects three types of intonational variety from the twenty Gp.I speakers.

(7) When the two complex dimensions are wholly independent (% diff. F, R and % diff. L, RF (RF being the fourth most frequent tone)), there is **no** regression of the sample from one axis to the other (Fig. B26). Nevertheless some of the groupings of speakers which are found in Fig. B25 are preserved, with some changes, in Fig. B26.

(8) There was little apparent discriminatory power in the distribution of nuclear selections in terms of tonicity (represented by the form class of the item carrying the nucleus), except a rather large difference in the % of nuclear common nouns (Q-high, $_{L}$T-low; cf. Fig. B27). The difference showed no subcategorisation of the $_{L}$T sample in terms of age, gender, or s-e class. Analysis which is not depicted indicates that this

difference between Q and $_{L}T$ is not lectal, but situational.

(9) A finding which is difficult to interpret is given by Fig. B28. Here apparently the tendency (amongst women) to realise R and FR on adverbs is high for young and old women and low for middle-aged women.

(10) Men tend to have a high frequency of nuclear adverbs which are F and a low frequency which are R and FR; and women conversely. (This is seen from plotting % nuclear advs. carrying R and FR against % carrying F, not here depicted.) But amongst men there is something of an age trend (young: higher F, lower R+FR; older: converse).

(11) Further patterns of dependency exist between other form-classes and terms from the system of tone, notably common nouns (n.s) and proper nouns (N.s). (Cf. Figs. B29, B30, B31, B32a, B32b.) Comparison of the distribution patterns of Fs

and Rs and Ls on n.s and N.s and adv.s shows that although the likenesses amongst the individual speakers (Gp.I), as one would expect, are different here (Fig. B32b) than they were on other dimensions (e.g. Fig. B26), there are interesting preservations. Such preservations encourage me to believe that the complex dimensions represented in Fig. B26 and in Fig. B32b are what I earlier have called, though in a slightly different context, **compatible criteria** (see above B(ii), p. 214).

(12) The most widely varying interaction (in Gp.I) between the systems of tone and pitch-range is Booster+Level (high for $_{L}$T, low for $_{NL}$V.s), though there are no marked interdependencies with social macro-variables except that women show less variation than men in respect of the numerical range of these co-occurrences.

(13) Fig. B34 shows that claims made in respect of Fig. B11 must be changed or reinterpreted. In particular:

(a) that lower values of the BF gradience are associated with $_{NL}V$.s must be rejected;

(b) that there is a parallelism between BF gradience and BR gradience (as in Fig. B11) seems to be false in Fig. B34;

(c) the ratio between BL/BF is roughly 90/60 in Fig. B11, but this is reversed in Fig. B34 (50/85);

(d) there seems to be a parallelism between BL and BR, representable as a mean difference of 11% (BL higher); I am unable to assign either a reasoned or a rational meaning to this;

(e) the value of BF does not predict the relative ordination of BL and BR.

D. Application of further speakers to the criteria: stability & chaos.

I now examine the grammatico-prosodic realisations of more speakers in order to try to resolve some of the uncertainties which have just been summarised.

Group II comprises 18 speakers - a random subset from the same sampling stratum as Gp. I - seven men and eleven women, distributed in age groups as indicated in Figure B35. Group III comprises 19 speakers from the same sampling stratum, nine men and ten women, distributed as indicated in Figure B36.

Even though these are three random groups (I, II, III) from a single sampling stratum of a random stratified sample, it is as well to reiterate two properties of random groups such that our interpretive expectations will not be drawn at too high a venture. Whilst

(a) there is a very large number of ways of drawing three theoretically perfectly random groups from a group of 57 speakers (i.e. the number of speakers in Groups I + II + III), nevertheless,

(b) it is equally perfectly possible for those three random groups

(i) **neither** to show homogeneous, or even similar, behaviours in response to some statistic,

(ii) **nor** to have homogeneous, or even similar, distributions of variable values or even of the variables themselves.

The numbering of the analyses which follow corresponds to the paragraphs in the summary of Section C, above.

C-Summary § (1)

The gross distribution of tones for the speakers in Gp. I (cf. Fig. B14) showed considerable differences from that for the Quirk et al. (1964) sample. In particular, it seemed as if some of the functions, or, at least, some of the distributions, of both F and of R tones in non-localised varieties were being taken over by RF and L tones in localised Tyneside varieties.

Figure B37 replots Gp. I against Gp. II and Gp. III and against Q. The inferences which I made concerning Fig. B14 are confirmed and strengthened by the distributions of tones in Gps. II and III. On overall average, Tyneside varieties, of the kind represented here, have roughly

20% fewer Fs,

10% fewer Rs,

10% more RFs, and

20% more Ls

than non-localised (Q) varieties.

In addition, contrasting the traces for Gps. I, II, and III, we may note that amongst localised Tyneside varieties, the lower the sum of R and FpR, the higher the sum of RF & L. (There is only one reversal, the value of RF in Gp. II being lower than it 'should' be.)

The stability of these differences seems to be well-established. They should therefore form a useful basis for determining criteria to discriminate between localised

Tyneside and non-localised varieties.

Of course Fig. B37 uses group sums (I, II, III) in its representation. But even if we represent the 57 individuals as individuals, by reducing quantitative to qualitative patterns, we can see that a finer coherent (?) structure of subgroups underlies Fig. B37. Figure B37a (two pages) gives, for each of the 57 informants, their ranking pattern on the six most frequent tonic frequencies (i.e. omitting RpF for all). The tones are in the fixed order of the Q frequencies, that is

F	R	F(pR)	FR	RF	L
1	2	3	4	5	6,

for all informants. There are **no** individuals in this sample of Tyneside speakers who share the Q pattern, as we might expect from what has preceded. The nearest is He 38, or perhaps, from different points of view, N 43, or Sco 15, or Rob 16, or Ak 21, or Wi 40. But in general we find, even amongst those Tynesiders whose most frequent tone is F, that the second most frequent tone is L, which is clearly very un-Q (rather than non-U !). The array of individual patterns, which though considerably variable is a good deal less so than it could be, is reduced somewhat in Figure B37b. There, I have grouped individuals by shared patterns which reflect and reiterate findings already alluded to. I have chosen to specify and examine only eleven patterns, some of which are composites ((7), (8) &(11)), though of course there are more distinct patterns than this in Fig.B37a. In Figure B37c I have re-presented the material of Fig. B37b as a subjectively drawn dendrogram - i.e. not according to any scaling of its basis - for those preferring 'visuals'.

First, we may distinguish type (iv) [pattern (11)] from the rest in terms of its being most like Q. Those speakers sharing pattern (11) have a low ranking of L associated with a fairly low ranking of R and a top ranking of F. I suggest that these are - in terms of prosodic and paralinguistic systems - speakers having transitional varieties of type a. That is, in terms of the various localised tonic patterns to be discussed below, these patterns are the ones one might expect to find amongst persons adopting strategies to reduce their localisedness.

The reduction of R frequency together with that of L is further evidence (see above, section C) for the existence of different types of Tyneside prosodic system resulting from a continuously varying interaction between the functions of Rs and the functions of Ls. [N]

[N] This discussion assumes that, other things being equal, the tonic functions to be realised in an informal interview will be more or less similar, as to range and frequency of occurrence, amongst different speakers from the same sampling stratum, such as these speakers are. In truth, I know of no positive evidence in favour of this assumption (general negative evidence is provided by the texts of Crystal & Davy (1969) from different situations). And anyway, for this kind of material, other things will very rarely be equal (B,C28).

The other patterns seem to me to be interpretable as comprising three broad types of localised Tyneside prosodic variety, variously composed of subtypes, as follows.

Type (i).

This consists of all the speakers who share patterns (1) - (5); the basis for the type being that all these patterns have a top ranking of L tone and a second ranking of F tone. The two subtypes may then be distinguished; subtype (a) in which the third ranked tone is RF [patterns (1), (2), (3)], and subtype (b) in which the third ranked tone is R [patterns (4), (5)].

Type (ii).

Here the characteristics are a high ranking of R and a high ranking of L. Type (ii) clearly shares more with type (ib) than with type (ia). All the members of type (ii) are women, three of them in the 31-40 age group.

Type (iii).

This type has a top ranked F (as in Q), a second ranked L and a low ranked R. This last characteristic makes type (iii) similar to type (ia) and dissimilar from types (ib) and (ii).

As we have seen, the transitional $_L$V.s [type (iv)] avoid each of these sets of defining characteristics by having a top ranked F and low rankings of both R and L.

C-Summary § (2)

It will be recalled that plotting age against the %

difference between Fs & Rs did not in fact discriminate localised Tyneside from non-localised varieties (Fig. B16), but showed a gender differentiation, and, amongst women, an age-trend. Figures B17, B18, and B19, explored these characteristics further but the only result of this further dissection was that I was unable to be confident about either the gender differentiation or the age trend.

Figure B38 plots Gp. II on age against the % difference between Fs and Rs. Here it is quite clear - at least from the evidence of Gp. II - that there is no gender differentiation. The mean values for the % difference are, crudely,

men: +19%,

women: +17%.

In addition, amongst the women of Gp. II there is also no indication of an age trend. Thus, in Fig. B38, both the continuous line encircling the men with high values on the dimension, and the dashed line purporting to indicate an age-trend amongst the women are in imitation of the same markings on Fig. B16. But

<u>unlike</u> that case, the number of exclusions required for such interpretations here (Gp. II), render them fantasies rather than possibilities.

A similar pattern of doubt, or rather denial, about both the gender differentiation and the age-trend is afforded by considering the plot of gp. III against age and the % difference between Fs & Rs in Figure B39. Here the complete invisibility of either pattern makes the interpretation of Fig. B16 seem, if not laughable, constituted by downright lies. (Means for both men and women are, crudely, +21%.)

The reason for wasting my own, and the reader's, time on these meaningless representations is an important one. I could have supressed Figs. B38 and B39, and limped along with the interpretation of Fig. B16, and the inconclusive speculations derived from it in Figs. B17, B18, and B19. But it seems to me important that one is both able, and permitted (not to say encouraged) to conclude concerning any segment of an investigation either:

(a) I do not **have** a plausible interpretation of these undoubtedly strong patterns of features/variables; or

(b) this data, in which I expected to find some patterns, doesn't **contain** any patterns (at least of the sort I expected or wanted) (B,C29).

There is, it seems to me, an unhealthy habit in academic, perhaps all, reporting which suppresses inconvenient data, excises unsupported hypotheses, and deletes partial or temporary contradictions of interpretation or theory. But the usefulness of another's thought is precisely the musings, the partly-baked ideas (Good et al.1962), the false starts, the blind alleys – in a phrase the reconsructibility of another's **motive** for, and **method** of, thinking about some problem. This, certainly, can never be got from results (having no exceptions) which exactly fit the predictions of a set of coherent hypotheses which, in turn, have been developed from a logically watertight theory. (Lakatos (1976: passim) has shown that none of it **ever** works this way in truth.) All of this must be doubly important in investigations such as the present one, in which no comparable work, or framework, exists.

Following Gleick (1988) and Hao (1984) I shall embrace the possibility that all apparently stable variations are created upon fractions of chaos each of which is necessarily unique (B,C30). There seem no very good reasons why such a possibility should be intrinsically less true of human behaviours, such as the patterns of realisation of prosodic and paralinguistic features, than of the formation of snowflakes, or of the indentation of coastlines, or of the number of petals in a single chrysanthemum bloom. (For further gesticulations in favour of such a view reconsider the sketches in Chs. I & 2 above.) Furthermore, precisely because the sample of material I am presenting here seems to have certain inherently self-contradictory or intractable characteristics in the face of the methods I am using to unravel its structure, **and** because, in spite of that appearance, I refuse to believe either that there is no structure in the material or that my methods are wholly inept, I shall try to suggest relationships between the stability and the chaos whenever I can see or imagine them. [N]

[N] The situation here sometimes seems almost as closed, and undecideable, as that outlined by Georges Perec (1988: xvi-xvii) for the jigsaw puzzler's relationship with the puzzlemaker:

"The art of jigsaw puzzling begins with wooden puzzles cut by hand, whose maker undertakes to ask himself all the questions the player will have to solve, and, instead of allowing chance to cover his tracks, aims to replace it with cunning, trickery, and subterfuge. All the elements occurring in the image to be reassembled . . . serve by design as points of departure for trails that lead to false information. The organised, coherent, structured signifying space of the picture is cut up not only into inert, formless elements, containing little information or signifying power, but also into falsified elements, carrying false information . . . Despite appearances, puzzling is not a solitary game: every move the puzzler makes, the puzzlemaker has made before; every piece the puzzler picks up, and picks up again, and studies and strokes, every combination he tries, and tries a second time, every blunder and every insight, each hope and each discouragement have all been designed, calculated, and decided by the other."

Or, put another way by Foucault (1983: 4):

"There is a worse kind of disorder than the *incongruous*, the linking together of things that are inappropriate; I mean the disorder in which a large number of possible orders glitter separately in the lawless and uncharted dimensions of the *heteroclite*; . . . in such a state, things are "laid", "placed", "arranged" in sites so very different from one another that it is impossible to find a common place beneath them all. *Utopias* afford consolation: although they have no real locality there is nevertheless a fantastic, untroubled region in which they are able to unfold . . . even though the road to them is chimerical. *Heterotopias* [e.g. Magritte, Borges, Foucault himself (JP)] are disturbing, probably because they secretly undermine language, because they make it impossible to name this *and* that, because they shatter or tangle common names, because they destroy syntax in advance, and not only the syntax with which we construct sentences but also that less apparent syntax which causes words and things (next to but also opposite one another) to "hang together". This is why utopias permit fable and discourse: they run with the very grain of language . . . ; heterotopias . . . dessicate speech, stop words in their tracks, contest the very possibility of language at its source; they dissolve our myths and sterilise the lyricism of our sentences."

Figure B40 plots age against the % difference between Fs and Rs for all three groups. (It is thus the sum of Figs. B16 (Gp I), B38 (Gp. II), and B39 (Gp. III).) As one would expect from the preceding remarks, very little structure is visible in this representation. All that can be said, and that with little interpretive force or content, is that the range of the % difference F/R is far larger amongst women (-33% to +45%) than it is amongst men (+2% to +50%), and that consequently the overall means for the genders are different (roughly, +12% for women, +22% for men). Nevertheless, it is important to emphasise that even this difference is brought about only by 5/28 of the women. A very weak age trend **may** be present, especially if we are willing to set aside the evidence for Ar (51-60, -25%); but the inferential framework within which I am operating here is so weak that we really have **no grounds** for rejecting her case. (And anyway all the 11 women below the age of 40 who have positive values militate strongly against the whole notion of an age-trend.) Perhaps the only significant characteristic represented here is that **no** men (29 cases) have negative values on the dimension (i.e. more Rs % than Fs %), whereas 5/28 women do have negative values.

What this might mean is an insoluble pancake (O'Brien 1967: 153). It is tempting to speculate that the prosodic behaviour of these women is in the opposite direction from that assigned to them by Trudgill (1972)(B,C31). It is clear, for instance, that the most reliable finding just referred to is not in the direction of any non-local prestige prosodic pattern (such as that which may be associated with the values for Q given in Fig. B37). And, on the other hand, it is clear that this behaviour is one which men have either retreated from, or one which they have not yet found the need or the basis to emulate. In either case, the women appear to be the guardians of a behaviour which is either <u>locally</u> prestigious, or not prestigious <u>at all</u> (B,C32).

<u>C-Summary § (3)</u>

Comparative examination of % Fs and % Rs against % diff. between Fs & Rs in Gps. II and III shows some agreement and some disagreement with the analysis of these dimensions in Gp. I. (The Gp. I analysis is given in Figs. B20a (for % Fs), B20b (for % Rs), and B20c (the two printed together).)

The % Fs are pictured in Figure B41 (Gp. II), Figure B43 (Gp. III), and Figure B45 (Gps. I, II, & III for comparison (i.e. Fig. B45 incorporates Fig.s B20a, B41, & B43 for convenience of reading)).

The % Rs are pictured in Figure B42 (Gp. II), Figure B44 (Gp. III), and Figure B46 (Gps. I, II, & III for comparison (i.e. Fig. B46 incorporates Fig.s B20b, B42, & B44 for convenience of reading)).

The most general agreement is provided simply by the equivalence of slope of the 'regression' lines of the three groups (I, II, III), for both Fs and Rs.

For Fs this is between 1 in 1 and 1 in 1.5 for all three groups. In other words, other things being equal, an increase of % Fs by 10% would predict - in the mouth of some newly sampled speaker - an increase in the positive % diff. between Fs & Rs of between 10% and 15%.

For Rs the slope is -1 in 2.8 for all three groups. In other words, other things being equal, a decrease of % Rs by 10% would predict an increase in the positive % diff. between Fs & Rs of around 28%.

Given that Gps. I, II, & III are three random sub-groups from the same stratum of a random stratified sample, these statistical regularities indicate that there is some stability of relation between this complex dimension (% diff. between Fs & Rs) and these simple dimensions (%Fs, %Rs). (Especially in the context of my remarks above (p. 270) about the possibility of obtaining inhomogeneous random groups from the same sampled stratum.) Clearly as a corollary of the above facts, just as in the analysis of % Fs and of %Rs against % diff. between Fs & Rs for Gp. I (Figs. B20a, B20b), the dependency between the distributions of Fs and Rs is not symmetrical in either Gp. II or in Gp. III. So much for the agreements.

The disagreements - between the analysis of Gp.I on the one hand and that of Gps. II & III on the other - are both more particular and more dramatic.

First, if we examine the gender distributions of Gps. II & III in Figs. B45 & B46, we find that they do not support the inference which I made in respect of Gp. I, namely that men tend to have higher values on the complex dimension than women. The trend associating higher s-e class with higher values on the complex dimension - which was weak enough in

Gp. I - also becomes invisible in Gps. II & III.

Secondly, we notice that, apart from Al, none of the members of Gps. II & III have negative values of the complex dimension. Any complexity of inference in the arguments above which was based on the positions of D, Ar, S, & Wa in Fig.s B20a & B20b must therefore be dubious, though not necessarily in error. The proportion of people having negative values on the dimension (more Rs than Fs) gradually drops from 4/20 (20%) in Gp. I, to 5/38 (13%) in Gps. I & II, to 5/57 (9%) overall. (A good illustration, though wryly admitted, of the variability of random groups from a single sampled stratum of a population !) In spite of this dramatic difference between the Gps., it is clear that the general regression of each group indicates the possibility of negative values on the complex dimension of the sort that we have in Gps. I & II.

Note in addition that though Gp. II shares with Gp. I a speaker having a positive value on the complex dimension of more than 45% , the highest positive value in Gp. III is 33%.

What I am emphasising here is that it is of no use to ask whether one or other of these groups of speakers is more typical than the other two, or, if you will, whether a rejuggled (therefore non-random) group of twenty of them might 'better' represent the rest.

The present state of understanding of intonational variation recalls the argument outlined by Bateson (1985: 34-6) in his illustration that science never proves, though it may improve and disprove, hypotheses. In the business of guessing the next term in a series, the assumption that one can predict what that term is rests upon both a preference for simplicity or parsimony, and a belief that there is an incomplete ordered sequence. (Thus: 2, 4, 6, 8, 10, 12, ? , usually yields 14. But if the series maker says "No the next term is 27", and then gives you 2, 4, 6, 8, 10, 12, 27, 2, 4, 6, 8, 10, 12, 27, 2, 4, 6, 8, 10, 12, 27, ?, you will probably say 2.) Bateson then writes:

Unfortunately (or perhaps fortunately), it is [the case] that the next fact is never available. All you have is the hope [my emphasis JP] of simplicity, and the next fact may always drive you to the next level of complexity. . . . For any sequence of numbers I can offer, there will always be a few ways of describing that sequence which will be simple, but there will be an infinite number of alternative ways not limited by the criterion of simplicity (1985: 36).

But this diffracts the perspicuity of the notion 'simplicity'. Would the answer 3, to the last series given, be one that was constrained by simplicity, or not ?

As far as the interpretation of the differences amongst Gps. I, II, and III on the complex dimension of Figs. B45 & B46 is concerned, I take Bateson's remarks to indicate that "the next fact" (or set of facts in the present case) has indeed moved us to another level of complexity.

And part of what has to be admitted in that case is that the kinds of intonational variety represented by the positions of the speakers D, Ar, S, Wa, and Al may be rare but they are not odd. [N]

[N] I have a trace memory that the distinction between rarity and oddity may be due to Michael Halliday, but I have no source for this trace, and the basis of my distinction may not be the same as his. **Rarity** is primarily a frequentist notion – the number of cases there have been, the number of tokens of a particular

type which have ever been seen, or manufactured, or recorded (see Ch. 3 above, & e.g. Preston: 1962). That having been said, there are clearly qualitative consequences of something being rare, in some such terms as its value, or its interestingness (Harrah 1963). **Oddity** is primarily a qualitative ascription arising from the structural abnormality or typological marginality of some object or phenomenon. Nonetheless there are quantitative implicatures in such ascriptions, since we often project our notions of structure or type onto hitherto unsampled populations in the form of expectations of frequency. In support of the secondariness of the qualitative component of rarity and of the quantitative component of oddity, one may cite the unexceptionableness of the contrasts between the possibilities of a thing being 'rare but not odd', 'rare and odd', 'odd but not rare', 'neither odd nor rare'.
[Michael Halliday (1990) confirms that the distinction in this sense **is** his, though "I didn't make the additional observation that you do, rightly - that each carries a secondary motif of the other"; and rightly points my defective memory to his paper 'Typology and the exotic' (in McIntosh & Halliday 1966: 166) as the source.]

May (1984) points out that simple non-linear equations of the type

$$X_{t1} = f\,(X_{t0})$$

(the value of some parameter *X* at a given time, *t1*, is a function of the same parameter at a previous time, *t0*) can rapidly lead to extremely complex dynamics in the system containing *X*. In other words such equations can be effectively insoluble unless such restrictive assumptions are made about f that the equation is made more or less linear.

The behaviour of %Fs and of %Rs on the complex dimension % diff. between Fs & Rs is clearly non-linear in May's sense. That is, we can write for Fs

$$F = f\,(F-R),$$

and for Rs

$$R = \Phi\,(F-R).$$

We know that f and Φ in any given system must be at least partially independent, since any given values for F and for R do not predict a value for (F-R), and since the variation of values for F over (F-R) does not predict the variation of values for R over (F-R).

C-Summary §§ (4) and (5)

Comparison of Figs. B20a, B20b, B21, B22, B23, and B24, above, showed that the part played by Ls in $_L$T varieties is very important and is also rather complex. Here, Figs. B21, B22, and B23 for Gp. I are compared with Figs. B47, B48, and B49 for Gp. II, and with Figs. B50, B51, and B52 for Gp. III.

The abscissa (x-axis) in all nine cases - x_{21}, x_{22}, x_{23}, x_{47}, x_{48}, x_{49}, x_{50}, x_{51}, x_{52} - is the same, notably, % difference between Ls and Rs (negative values signify more Rs than Ls, positive ones more Ls than Rs). The ordinate (y-axis) in B21, B47, and B50 is %Ls, in B22, B48, and B51 is %Rs, and in B23, B49, and B52 is %Fs. [N]

[N] ! ! ! The ordinates of Figs. B21, B22, and B23 are half the scale of those in B47, B48, B49, B50, B51, and B52. I apologise to the reader for this additional burden.

If we examine both the 'regression' slopes and the rank preservation of speakers amongst these depictions we find both stability and chaos. Although the distribution of Ls is exceedingly complex and indicates that Ls have different functions and distributions in different $_{L}T$ varieties, and although the patterns of Ls, Rs, and Fs are different in the three random subgroups (I, II, III), there is a degree of similarity between the three groups sufficient for us to have some confidence in the stability and utility of the dimensions we are using.

<u>Slope</u>

Let us first consider the slopes. [N]

[N] Notice that the dispersion of the speakers in Figs. B21, B22, B47, B48, B50, and B51 is more 'orderly' or 'linear' than it is in Figs. B23, B49, and B52. This is because, though we cannot specify the functions γ, τ in

$L = \gamma(L-R)$ (B21, B47, B50),

$R = \tau(L-R)$ (B22, B48, B51), $[\gamma \neq \tau]$,

for the sorts of reasons adduced above, any predictions of the values of γ or τ will always be better than a prediction of the value of θ in

$F = \theta(L-R)$,

since F is totally unrelated to that which θ dominates (B,C33).

The slopes for %Ls on % diff. between Ls and Rs are:

Gp. I 1 in 1.3

Gp. II 1 in 1.5

Gp. III 1 in 1.3 (mean 1.4),

and the points at which the traces in the three cases cut the y-axis are between 12% and 17% Ls. This is an impressive level of agreement and encourages us to look at the other cases.

The slopes for %Rs on % diff. between Ls and Rs are

Gp. I -1 in 1.4

Gp. II -1 in 3.25

Gp. III -1 in 4.25 (mean 3.0),

and the cutting points on the y-axis are between 15% and 20% Rs.

The slopes for %Fs on % diff. between Ls & Rs are equivocal, or double, depending on how one is looking at them, as follows:

(a) slopes more or less tangential to a circle with Q at its centre (i.e. uphill slopes) -

Gp. I 1 in 1.3 (this is only clear because of D, Ar, S, & T),

Gp. II 1 in 2.5 (this is not obvious [N]),

Gp. III 1 in 1.5 (this is very unclear), (mean 1.8);

(b) slopes more or less on a radius of a circle with Q as its centre (i.e. downhill slopes) -

Gp. I -1 in 1.0 (this is unclear),

Gp. II -1 in 3.5 (this is not obvious [N]),

Gp. III -1 in 2.8 (clear), (mean 2.4).

[N] The problem here is that the slopes for Gp. II in Fig. B49 could just as easily be specified as **zero** (i.e. horizontal). The form in which they are expressed above represents a desire for similarity or compatibility between these slopes and others in the series. No help towards such a goal is forthcoming from either Gp. I, in which the positive slope is only established on the varieties of D, Ar, S, & T (who we have already said are rare), and in which the negative slope could be anything between -1 in 1.0 and -1 in 0.1,
or Gp.III in which the positive slope is extremely tenuous and in which the clear negative slope is -1 in 2.8.

The slopes for %Fs on % diffs. between Ls & Rs (i.e. Figs. B23, B49, B52) are so chaotic that there is little point in even trying to consider cutting points.

In spite of the range of differences between Gp. I, Gp. II, and Gp. III in respect of %Ls, %Rs, and %Fs on the complex dimension (% diff. between Ls & Rs), and in spite of our inability to predict γ, τ, or θ with any likelihood of success, it is clear that the complex dimension (%diff. between Ls & Rs) is stable for new (i.e. Gps. II & III) samples of speakers.

Even the chaos represented by the slopes described for Figs. B23, B49, and B52 confirms the main inference made above solely for Gp. I. (Namely, that when explicit dimensional power is given to Ls, localised varieties are represented on two distinct trends, one more or less a radius and one more or less a tangent of a circle centred on Q.) In fact, when all three groups are plotted by %Fs on % diff. between Ls & Rs (i.e. the sum of Figs. B23, B49, & B52, not depicted here) this pattern is repeated, with a strong negative (downhill) slope of -1 in 3.4 and a weaker, positive, slope of 1 in 2.

Rank preservation

We are here comparing the relative rankings of speakers on the y-axes of successive pairs of graphs of the same speakers:

for Gp. I : Figs. B21, B22, B23;

for Gp. II : Figs. B47, B48, B49;

for Gp. III : Figs. B50, B51, B52;

respectively, %Ls, %Rs, %Fs.

The following matrices show the number of speakers, for each pair of graphs, who are displaced by each of the possible number of ranks. (See p.236ff., Section C, for the method of computation.)

First, we compare the ranking on %Ls with that on %Rs. (Of the ordinates, y_{21}, y_{47}, y_{50} are % Ls, y_{22}, y_{48}, y_{51} are % Rs.)

No. ranks sp. displaced by (pos & neg):	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
No. sps. so displaced comparing																			
y_{21}/y_{22}	1	1	5	1	2	2	-	1	2	1	-	1	-	1	-	-	2	-	-
y_{47}/y_{48}	2	2	2	1	2	2	-	-	1	1	2	-	1	-	1	1	-	-	-
y_{50}/y_{51}	2	1	3	-	2	1	-	-	1	1	-	3	1	1	-	2	-	1	-

Secondly, we compare the ranking on %Ls with that on %Fs. (Of the ordinates, y_{21}, y_{47}, y_{50} are % Ls, y_{23}, y_{49}, y_{52} are %Fs.)

No. ranks sp. displaced by (pos & neg):	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
No. sps. so displaced comparing																			
y_{21}/y_{23}	2	1	5	-	-	-	1	1	2	1	1	3	-	1	-	1	-	-	1
y_{47}/y_{49}	-	1	2	-	1	2	2	1	-	3	2	1	-	1	1	1	-	-	-
y_{50}/y_{52}	-	1	1	1	1	1	1	-	2	1	4	-	1	2	-	1	-	1	1

Thirdly, we compare the ranking on %Rs with that on %Fs. (Of the ordinates, y_{22}, y_{48}, y_{51} are % Rs, y_{23}, y_{49}, y_{52} are %Fs.)

No. ranks sp. displaced by (pos & neg):	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
No. sps. so displaced comparing:																			
y_{22}/y_{23}	3	2	6	2	1	-	2	2	1	-	1	-	-	-	-	-	-	-	-
y_{48}/y_{49}	1	1	1	2	2	1	-	3	-	2	-	-	-	2	2	-	-	1	-
y_{51}/y_{52}	2	3	3	3	1	-	1	1	-	-	1	1	2	-	-	-	1	-	-

Lastly, we may get a qualitative feel for the functions γ, τ, and θ across the three Gps. (I,II,III), by constructing a table which shows the proportion of speakers in each group, for each comparison, which suffers five or fewer displacements and the proportion which suffers eleven or more displacements. (Recall that the functions γ, τ, and θ are determinants in the equations:

$L = \gamma(L-R)$; $R = \tau(L-R)$, and $F = \theta(L-R)$.)

I am not here suggesting that we can find values or trajectories for these functions (which is to be attempted in Pellowe (in preparation, a)). They are simply shorthand labels emphasising the complexity of the outcomes on our simple & complex dimensions.

		Proportion of sps. displaced by ≤ 5 ranks	Proportion of sps. displaced by ≥11 ranks
Comparing %Ls & %Rs			
in	Gp. I	12/20 (60%)	4/20 (20%)
	Gp. II	11/18 (61%)	3/18 (17%)
	Gp. III	9/19 (47%)	8/19 (42%)
Comparing %Ls & %Fs			
in	Gp. I	8/20 (40%)	6/20 (30%)
	Gp. II	6/18 (33%)	4/18 (22%)
	Gp. III	5/19 (26%)	6/19 (32%)
Comparing %Rs & %Fs			
in	Gp. I	14/20 (70%)	0/20 (0%)
	Gp. II	8/18 (44%)	5/18 (28%)
	Gp. III	12/19 (63%)	4/19 (21%)

Of this table I note the following.

(a) The values of γ (in $\bar{L} = \gamma(L-R)$) and of τ (in $R = \tau(L-R)$) are in conflict in some way in the case of Gp. III, since 42% of its speakers are displaced by eleven or more ranks (and only 47% by five or fewer ranks).

(b) Similarly, the values of τ (in $R = \tau(L-R)$) and θ (in $F = \theta(L-R)$) are in conflict in some way in the case of Gp. II, since only 44% of its speakers are displaced by five or fewer ranks.

(c) Finally, the interactions between γ and θ are the least regular, since they lead, for all three groups to low proportions of speakers displaced by five or fewer ranks and to high proportions of speakers displaced be eleven or more ranks.

In general, what I wish to suggest from this examination of Figs. B21, B22, B23, B47, B48, B49, B50, B51, and B52 is that the characteristic slopes and the rank preservations in the various cases argue <u>both</u>

(i) that the dimensions as defined and exploited are demonstrating their stability and compatibility (see fn. above, Section B(ii), p. 214ff.) in the face of new data; <u>and</u>,

(ii) that there is a necessary indeterminacy which limits our capacity to represent variant intonational systems <u>in toto</u>. This intrinsic unpredictability arises from the cumulative effects of the non-linear relations between terms in variant underlying prosodic systems.

C-Summary §§ (12) and (13)

In my application of further speakers to the criteria I turn, finally, to the problems of tone + booster which were first raised in discussion of Figs. B11 and B34. [N]

[N] The reader may be relieved to know that I shall not be dealing in this thesis with the Gp. II and Gp. III values in respect of §§ (6), (7), (8), (9), (10), and (11) of the summary of Section C, above (p. 263 ff.). They are to be examined in Pellowe (in preparation, a).

Fig. B34 (boosted Fs, %Fs; boosted Ls, %Ls; boosted Rs, %Rs ranked on BF, for Gp. I) showed that regularities noted in, or inferences based on, Fig. B11 had to be changed or rejected. In particular:

(a) that lower values of BF gradience are associated with $_{NL}Vs$ must be rejected;

(b) that there is parallelism between BF gradience and BR gradience seems to be false;

(c) the BL/BF ratio is 90/60 in Fig. B11 but is more or less reversed (50/85) in Fig. B34;

(d) there seems to be a parallelism between BL and BR (a mean difference of 11% (BL higher));

(e) the value of BF does not predict the relative ordination of BL and BR.

Figure B53 plots boosted Fs, %Fs; boosted Ls, %Ls; boosted Rs, %Rs, ranked on BL, for all three Gps. (I, II, III).[N]

[N] Direct comparison between Figs. B34 and B53 is not easy (but is possible) since they are ranked on different parameters. Fig. B34 is ranked on values of BF, Fig. B53 is ranked on values of BL. Since Fig.B34 is 'contained in' Fig. B53, but in a different form, direct comparison is not necessary. All we need to do is to examine Fig. B53 in respect of the conclusions arrived at on the basis of Fig. B34.

(a') In agreement with the facts of Fig. B34, low BF does <u>not</u> indicate $_{NL}$Vs, because

(i) those speakers with values of < 60% BF are known on non-prosodic grounds to be $_{L}$T speakers (<u>An</u> (12), <u>S</u> (41), <u>Dru</u> (56), <u>Boy</u> (57), <u>Mar</u> (59));

[but the implicature here assumes that the solidarity principle holds which, with Bazell's (1949, 1966, 1977b) help, I have already cast doubt upon above (Ch.4)];

(ii) those varieties which we know on non-prosodic grounds to be less- or transitional- localised have BF values which are not particularly low (N (94), Ch (91), Mc (86), Su (79));

[same remark about the implicature];

(iii) varieties which in other respects have been referred to as maximally localised show a considerable range of BF values (D (97), Ar (73), S (41));

[same remark about the implicature].

(b') The value of BF for any particular speaker does not predict BR (and conversely) and the value of BL can predict neither BF nor BR in particular cases (but see (d) below).

(c') The BL/BF ratio ranges from 67/84, through 54/74 to 25/56, with let's say roughly a mean value of 23%.

(d') The parallelism between BL and BR permits one to predict that in general, for a given value of BL, the value of BR will be 22% less. Such a prediction may, in particular cases, be in error by anything up to +40% or -20%.

(e') Though it is true that the value of BF does not predict the relative ordination of BL and BR, there is a remarkable stability of ordination overall. After inspection of Fig. B53, we would expect, in general, to find the values of BF, BL, and BR for a newly sampled Tyneside speaker to satisfy the inequalities -

$$BF > BL > BR.$$

The patterns of exception are interesting

		BL > BF	BR > BL
Gp. I	(Fig. B34)	2/20 (10%)	5/20 (25%)
Gps. I, II, III	(Fig. B53)	2/57† (3%)	14/57 (24%)

(† the only case of BR>BF (i.e. informant An) is also one of the two cases of BL>BF).

This makes it look as if BR>BL is only a relatively rare category, whereas, given the increase in the number of cases, BL>BF (and, by implication, BR>BF) may be a category of some oddity. (As I remarked in my note above, oddity certainly has quantitative associations, here in addition, and perhaps paradoxically, I am implying that oddity may on occasion have quantitative origins.)

On the basis of the present material alone, it is impossible to assign an extralinguistic meaning to these regularities. But what Fig. B53 does tells us is that the interactions between the system of tone and the (generalised) terms from the simple system of pitch range are stable enough to be worthy of further use in investigations of the prosodic variation of lects.

E. Establishing prosodic and paralinguistic criteria experimentally.

Here I present an alternative way - it is possible, even, that we shall witness a modest amount of convergence - of determining the variability of intonational systems. But it certainly could not be guaranteed that the two methods would ever be able to produce identical results. Nevertheless, at this stage in my attempts to find a way of even expressing that which we all naively and anecdotally know about intonational variation, any way forward is something of a surprise and an adventure. An additional bonus accrues from the fact that, compared to the costly business of collecting and analysing survey data, this method is relatively cheap in terms of person-hours. (What follows draws on Pellowe 1980.)

The method however, is much more than merely a value-added luxury for dealing with intonational variation. Precisely because it is experimental and not 'real' language use, the spoken output of the experimental volunteers is less likely to vary in unwanted ways than interview material. (This almost certainly begs the question of whether anyone can ever know exactly what all the unwanted manners of varying are, and why (B,C34).)

But to give substance to this, let me re-examine some analysis we have seen before:

Consider again Figure B14 - the gross % distribution of tones for Gp. I - of which I suggested that the differences between the T(yneside) and Q(uirk) traces could be thought of as being diagnostic of two sets of varieties. (The critical differences are the relative distributions of F, RF and L tones.)

What is completely unclear however, is exactly what it is that is being diagnosed. In this respect there seem to be two polar possibilities.

Either, (1) the differences in Fig. B14 are diagnostic of the difference between localised Tyneside and non-localised (educated) varieties. (In the distinctions offered by Fig. B1, the differences are taken to diagnose varietal variability.)

Or, (2) the differences in Fig. B14 are diagnostic of the difference between varieties of speech appropriate to informal interviews (T), and varieties of speech appropriate to broadcast panel discussion (Q). (In the distinctions offered by Fig. B1, the differences are taken to diagnose realisational variability.) The first possibility would encourage the view that the differences arose from different underlying systems. The second would attribute the differences to different communicative functions, and would be neutral as to sameness or difference of underlying system.

Clearly a mixture of these two possibilities (1 & 2 above) is not only possible but, given the ineffable nature of things, likely. For ease of argument at present I shall ignore the mixture problem, but I will return to it below.

Comparative analyses of pieces of spoken language having widely differing communicative functions have been published by Crystal and Davy (1969). Their results indicate that the absolute frequency of, and the co-occurrence rates between, prosodic and paralinguistic features differ considerably for such different communicative functions.

Now, given that one takes the findings of Crystal & Davy seriously, and given that one accepts the difference between T and Q in fig. B 14, one may **test** for one or the other of the possibilities above (1 & 2), according to the following program and its implied argument.

(i) Define some dimension based on the diagnostic differences of Fig. B14. Let us take criterion (1) of Fig. B12 - the difference between the frequency of occurrence of falling tones and that of rising tones. This dimension, according to Fig. B14, will have high values for Q-type varieties (whether they turn out to be non-localised or panel discussion) and low values for T-type varieties (whether they turn out to be Tyneside-localised or interview).

(ii) Next, plot values of this dimension against a non-linguistic axis, such as age. (This will have to be for the T sample only since I have no relevant data for the Q sample.)

(iii) **Then,** if the differences in Fig. B14 had nothing to do with reflecting the differences between localised-Tyneside varieties and non-localised varieties (possibility (1) above), but had only to do with the differences between interview varieties and panel-discussion varieties (possibility (2) above), then we would expect such a picture to show a random, or at least unstructured distribution.

That is, it is extremely difficult to imagine what might be the role of age as an independent variable showing the structure of tonic distributional differences between discussing things on a panel and being an interviewee in your own home.

On the other hand the more that the differences in Fig. B14 have to do with the differences between non-localised varieties and localised-Tyneside varieties (alternatively the more they have to do with different preferential prosodic patterns - what I am calling different underlying or lectal systems), then the more we might expect structured differentiation in the picture we have suggested (age by % difference between falls and rises - cf. Fig. B16).

This argument is based on what I deem to be an uncontroversial reversal of normal sociolinguistic practice:

rather than saying: 'here is a linguistic structure the fact of whose variation I know about - what are its extralinguistic correlates ?', one can say: 'here is an extralinguistic correlate of much linguistic variation - if it shows structure in the variation of some prosodic system, we can assume that that prosodic system is varying in the required sense -

that is, as to lect, rather than as to communicative function.

Detailed interpretations of Figures, such as B14 and B16, showing the distribution of various prosodic features and items have been presented above. Here I do not want to repeat them, but to reiterate a general claim.

I think that one can be fairly certain that this method of distinguishing between truly equivalent and speciously equivalent formally identical prosodic features - a method of closely examining a whole series of partially overlapping depictions - gives a reasonable representation of the lectal variability in prosodic systems.

However, the reason why one can't be more than fairly certain of the reasonableness of this representation has to do with the sameness of the expressive resources for two kinds of prosodic variability which need to be distinguished. Variability which is due to difference of communicative function rests on the **same** basic set of distinctive linguistic features (different types and levels of patterning of speed, loudness and pitch) as variability which is due to lectal differences.

That is, it is just conceivable that a not insignificant fraction of the regularity of differences amongst these Tynesiders has not to do with different underlying systems, but has to do with differential interpretations of the nature of the communicative situation that they were in (ranging, perhaps, from 'torture', through 'interrogation' and 'cross-examination', to 'inconsequential chit-chat'). [N]

[N] In order to explain some of the findings of Pellowe & Jones (1977) and the interpretations of Figs. B20a, B20b, B20c on the one hand & B21, B22, B23, B23a on the other (Section C, above), it would have to be the case that these differential interpretations of the nature of the communicative situation were <u>ordered</u> with respect to one, or several, independent ('social') variables (??personality type, ??age, ??gender . . .). I do not entirely reject this possibility, but neither do I know of any evidence in its support.

In order to try to partial out this possibility it is necessary to try to obtain data in which the possible contribution of this kind of variability (realisational variability) is reduced to a minimum. To do this I have imitated an experiment of Quirk's and Crystal's (1966). [N]

[N] I am very grateful to Randolph Quirk for providing me with a copy of the tape of their stimulus material.
My version of Quirk's and Crystal's experiment is not, strictly, a replication for reasons which will become clear below. I did not use the same stimulus material and I did not have the same purpose. As will become clear, I believe that their purpose may not be achievable before my own, even though the method I use is due to them.

In their (1966) work Quirk and Crystal (QC) aimed to determine the degrees of contrastivity attaching to the various contributory features of tone units in educated spoken British English. How contrastive is the stretch of utterance incorporated in any given tone unit ? How contrastive is the location of the tonic nucleus ? How contrastive is any particular tone in respect of all others ? How contrastive is the initiation of a tone unit (onset) ? How contrastive are stress and pause ? [N]

[N] It is another rough measure of our ignorance about prosodic and paralinguistic systems that these questions do not sound as strangely elementary as, for instance, 'How contrastive is voicedness in segmental phonology ?' or 'How contrastive is [+female] in the subset of the lexicon labelled [+animate] ?'

Quirk & Crystal give answers to these questions by using the following method (1966: 359-361). A recording was made of a realisation of the sentence

A few minutes later the fellow came. He walked up slowly and said, 'Oh! It wasn't you that I wanted'.

The recording was played right through to each informant. It was then played in two halves (with the break after *came*), the informants being asked to repeat each half after it was played.

There were 46 informants. Transcriptions were made of the model sentence and of each informant's version of it. Scrutiny of the transcriptions of the informants, as against that of the model, revealed differing degrees of agreement and disagreement with various features of the model sentence.
It was argued by QC that high levels of agreement are correlated with high contrastivity of the feature concerned. Nevertheless, as Quirk & Crystal point out (1966: 361-2), the contrastivity of a particular substantial feature is not a fixed quantum but varies in respect of its prosodic, paralinguistic and syntactic environments. In spite of this problem they found that stable ranges of contrastivity could be established for features of different, hierarchically related, systems. Particular results of theirs will be discussed as they impinge on the results of the present work. One of the sources of variability as to agreement and disagreement which Quirk & Crystal did <u>not</u> consider in their paper is precisely the one which motivates the present work.

One might argue that insofar as a given substantial feature in the model is **not** agreed upon (i.e. repeated) by an informant, or a group of informants, just so far we may attribute a different preferential pattern of realisation to those informants in respect of that type of feature, in that frame, in that environment (B,C35).

Instead of viewing degrees of agreement and disagreement between informants and target models as indices <u>only</u> of the contrastiveness of **features in the models**, we can view them as measures of the varying salience of the features in the models in respect of the **intonational systems of the individual informants.**

In this view there are thus two possible sources for agreements and disgreements with particular features of a model. These are:

(a) varying degrees of contrastivity of the prosodic features which are present in the model sentence;

(b) variant intonational systems in the response repertoires of different informants.

Clearly, insofar as the range of differences amongst informants' own underlying systems increases , so far also different informants may be expected to have wideningly different agreement rates with some particular feature of a model (B,C36). In other words (a) and (b) above may be expected to interact. We can accept that rates of agreement and disagreement with features of a model have **only** to do with the degrees of contrastivity of those features in those contexts under the strict assumption that the intonational systems of the speaker of the model and the intonational systems of the informants responding to the model are the same (assumption 1)(B,C37).

In the work now to be discussed this assumption is deliberately discarded. In its stead I make the assumption (assumption 2) that

the range and type of agreements and disagreements which a given informant characteristically has in respect of the various features in several models is a reflection of her preferential realisation patterns as a speaker under normal circumstances, and that these preferential realisation patterns are reflections

<u>of her underlying intonational systems</u> (B,C38).

I advance, under assumption 2, four crude hypotheses:

(H1) Informants whose patterns of agreement and disagreement with models are similar, have similar intonational systems (B,C39);

(H2) Informants whose patterns of agreement and disagreement with models are different have different intonational systems (B,C40);

(H3) Informants showing a generally high level of agreement with some model or models have similar intonational systems to those underlying the model or models (B,C41);

(H4) Informants showing a generally high level of disagreement with some model or models have dissimilar intonational systems from those underlying the model or models (B,C42).

The crudity of these hypotheses, as we shall see, results from their failure to take account of the interaction between the lectal variability of intonational systems and the varying degrees of contrastivity in some particular system. However the substance of the final remark in the penultimate footnote (p.315) will now be clear. If assumption 2 above is reasonable, then the types of lectal variability in intonational systems have to be established before one can determine a suitable sample of speakers in order to satisfy the stricture in assumption 1 which I claim is necessary for the prosecution of QC's purpose.

Model sentences and their variation.

In the Quirk & Crystal experiment (1966), there was a single model sentence (spoken by DC). Because of the different purpose which I have here (cf. assumption 2 above), I needed a range of different models to which the same group of informants could respond (cf. H1 to H4 above). At the same time I wanted to be able to compare the results I obtained - even under the different assumption - with those of QC. Consequently I needed to generate different models of the same sentence as theirs, and models having some chance of being realised in comparable units, with comparable occurrences of slots for the realisation of particular substantial features. In other words, I wanted to obtain models which could reflect the variant underlying systems of their speakers, but at the same time be comparable with the QC model. [N]

[N] The realisation of the model in QC seems to me a not unmarked form of spoken English. Firstly, given the high speed of the realisation, the tonic realisations are too well defined - in my view wider than the speed would predict for that variety of English. Secondly, when I played it to an audience having varying degrees of linguistic expertise, it elicited judgemental labels (of variety) such as: 'for learners of English as a foreign language' 'radio acting register'. There may be things wrong with my own models , but I have not found that they elicited remarks either as definite, or as marked, as these.

To achieve these ends the following procedure was adopted. Nineteen volunteer students, whose segmental phonological realisations indicated that they came from widely differing speech communities in the UK, made recordings of the model sentence under language laboratory conditions. [N]

[N] The localisation of the phonology of the students' varieties had been independently checked from a biographical questionnaire they had filled in some weeks before.

I shall henceforth refer to these informants as 'modellers'. The modellers were provided with a 3-slip booklet which they were asked not to open until they were told to. The first slip was blank. The second and third slips bore the parts of the same model sentence as that of Quirk and Crystal (1966). The layout of the sentence on the slips was as follows:

I	A few minutes later,
I	the fellow came;

|	he walked up s l o w l y and said "Oh !,
|	it wasn't <u>you</u> that I <u>wanted</u> ".

Modellers were told that the second and third slips bore parts of a sentence and received the following instructions (which were repeated):

> 'When you are signalled to begin, will you turn back the blank slip and speak what you find on the other slips as naturally as you can, paying due attention to how it is laid out on the slips.'

After this recording had been made the modellers were instructed as follows:

> 'Before, during, or after your speaking of the sentence, you may have wondered what, exactly, was meant by the phrase 'paying due attention to how it is laid out on the slip'. I would like you to think carefully for a minute about the layout on the slips. You may feel that, on reflection, you would not have said the sentence how you <u>did</u> say it if you had had time to think about it. Even if you feel that your first speaking of the sentence was <u>right</u> for the layout, I would like you to make a second recording'.

After a minute's pause the initial instruction was issued again, and the modellers recorded a second version.

There are of course arguments against the use of written stimuli for such experiments as this. Two points are worth making:

(a) given that one wants to obtain fairly exactly comparable material, for the reasons already stated, there seems no alternative method of obtaining the model realisations;

(b) judges to whom I have played my models have not opined that the sentence was being read.

The function of the layout on the slips was to encourage, but not to force, a parallelism of occurrence of tone units and hence, of realisation slots, between my models and that of QC. The TUs in QC's model (tonic syllables underlined) are:

A few minutes <u>lat</u>er,

the fellow <u>came,</u>

he walked up <u>slowly,</u>

and said <u>Oh,</u>

it wasn't <u>you</u> that I <u>wanted</u>.

The relationship between these TUs and the layout on the slips may easily be seen. [N]

[N] There appears to be only one major discrepancy between my models and that of Quirk & Crystal which may be attributable to the written stimulus and its layout. That is the not uncommon occurrence amongst my models of both onset and nucleus, or simply onset, on said. This may well be caused by the "shift" effect of the quotes preceding Oh! (and of the upper case letter 'O' itself).
In the QC model and said occurs as the pre-onset part of the TU whose onset and nucleus both occur at Oh!. Notice however that 32 of their 46 informants shifted the onset from Oh to said.

These model sentences were played to informants who were asked to repeat them. Before discussing that process however, I need to give some indication of the range of variations amongst the models.

In the rest of what follows concerning this imitation of QC, I shall only be dealing with the first fifteen of the thirty-eight models (i.e. 19 x 2) which were produced. These fifteen models are composed of 2 models each from 7 modellers, and one model from an eighth. [N]

[N] The remaining 23 models - and their imitations - are to be tested independently at a later date against the findings to be presented now.

Figure B54 shows the range of tonic variation in terms of location of nucleus and direction of pitch movement for pairs of models produced by 7 speakers according to the method specified above. The speakers (H, W, P, etc.) are ranked from left to right in terms of 'increasing northernness' on the basis of a crude overall assessment of segmental phonology. The tonicity of any particular nucleus in the body of the table can be read off from the sentence running down the side. Nuclei are represented according to a ten-point typology:

rise (R), fall (F), level (L), rise-fall (RF), fall-rise (FR), rise-fall-rise (RFR), fall of fall-plus-rise (F(FpR)), rise of fall-plus-rise (R(FpR)), fall of rise-plus-fall (F(RpF)), rise of rise-plus-fall (R(RpF)).

The tonicity and nuclear pitch movements in the QC model are given in the right-most column. The norm ('mean') tonicity and nuclear pitch movement, for this group of models, is given in the right-most but one column. H1 is modeller H's first model, H2 her second. Three indexes are given below each modeller's second model:

(i) the total number of changes made from that modeller's first model;

(ii) the number of qualitatively different changes made by that modeller;

(iii) the number of different changes which were changes to more complex tones.

There is certainly variation amongst these models, but the most important thing to be noted is the generally good agreement with the QC model in terms of the abstractions of tonality (tone unit domain), and tonicity (location of the nucleus). The insertion of **additional** tone units (with nuclei on few, said and wasn't) seems to be favoured by the less-northern modellers. Of those having said as a nucleus, the selection of level rather than rise as its tone seems to be less-northern. More-northern modellers seem to prefer a simple tonic movement for wanted than do less-northern modellers. The most changed nucleus (as between the first and second versions of all modellers) was that on slowly. This is probably because - as some modellers said after the recording session - the typography of slowly on the slip was one of the features not noticed until pause for thought was given. There seems to be a slight tendency for more-northern modellers to change fewer nuclei between the first and second versions than less-northern modellers. Of the different changes which are made by modellers the more-northern modellers seem to change to more complex nuclei less frequently.

Figure B55 presents similar information for sixteen models, (i.e. fifteen of my models plus QC's), but in quantitative form. The Fig. gives the frequency of selection of a given tone for a given nuclear syllable in terms of the complying number of models. The underlined number for any particular nuclear syllable (row) indicates that the same tone is the one used in the QC model. The two elements of compound tones are linked by a dotted line. The two columns on the right give the typicality (√) or atypicality (x) of the presence (left column) or absence (right column) of the features in the QC model, in comparison with the numbers provided by my models.

We may note firstly the considerable variability of tonic selection on later. Quirk and Crystal attribute the performance of their imitators - twelve out of forty six of whom did not repeat a TU for the stretch a few minutes later - to the low predictability of a TU in front-placed verbless adjuncts. Yet all fifteen of my modellers **did** select that stretch as a TU. Could the manner of presenting the stimulus, with later at the end of a line, completely overcome such low predictability ? (See also below under 'Imitation variation'.)

If low predictability is to be understood as equatable with high information value (species Shannon), then it is surprising that such a wide range of tonic selections is possible. If low predictability is . to be understood as equatable with low information value (marginality), then it is surprising that so many modellers choose to delimit it as a tone unit (B,C43).

The rise realised on later by the QC model is atypical in terms of my models in which fall-rises are used by six out of fifteen, falls by four out of fifteen. Note also that though the absence of tonicity on few in the QC model is typical, five out of fifteen (33%) of my models (three speakers) do make it tonic. This is interesting in terms of a conclusion of QC's - two of whose forty six (4%) imitators made few tonic - to the effect that tonicity on pre-heads in NPs is rare.

I have already mentioned reasons why my models render QC's model atypical in respect of non-tonic said, which when it is nuclear clearly affects the tonic distribution for slowly in terms of the possibility of compound tones, e.g. fall(slowly)+rise(said). Of greater significance are the tonic selections for you and wanted. These distributions are such as to make a good deal less certain QC's claim that, in the TU wasn't you that I wanted, with onset on wanted, and fall+rise realised on you, wanted, the items you and wanted "are correlative and interdependent, forming a complex without the possibility of intervening onset" (1966: 361).[N]

[N] In addition, consider my model 38 (from modeller 19 - not used here) which has the following TU structure:
/F Oh * it /wasn't FR you * that / F (wide) I * / RF wanted *

Figure B56 shows the degrees of agreement amongst my 15 models, and between them and the QC model, in respect of tonicity, onset and pause. The scale on the left represents the number of my 15 models having a particular feature at a particular location in the model; the QC location of some feature is indicated by x attached to the relevant part of the trace for my models. The highest levels of divergence, for all three features, between QC and these models, cluster around the stretch and said Oh.
Out of 15, 10 models have onset on said, 11 have nucleus on said, 9 have pause between said and Oh. None of these features occur in these locations in QC. But these distributions are of exactly the same type as those in my model which **do** match the QC model. For

example, 11 out 15 have onset on Oh, 15 have nucleus on Oh, 10 have pause between Oh and it.

Minority divergences of tonicity in my 15 models are represented by: 4 few; 6 fellow; 3 walked; 6 wasn't. Other divergences, of onset, are predicted by these facts. The pause after later in 7/15 of my models is probably a direct effect of the instructions to modellers, and of the layout on the slips - later was at a line-end. We cannot compare the pause between came and he in my models with that which occurs in QC, since my modellers had to turn a slip at that point, i.e. they had to pause of necessity. [N]

[N] Pauses for one modeller (Hu) are not included in Fig. B56. In her first version (Hu1), she had pauses between up and slowly, and between it and wasn't. In her second version (Hu2) she had pauses between slowly and and, and between I and wanted. The consistency of the use of pauses in connection with the adverbial, and with one of the verbs of the cleft, in spite of their mobility in the two versions, indicates that their presence may not be erroneous (for her).

Figure B57 shows the variability of the 15 models in respect of features from the simple and complex systems of tempo (Crystal &

Quirk 1964; Crystal 1969). The first and second versions of each modeller are ranked down the left hand side of the figure in decreasing order of speed, roughly speaking. [N]

[N] Such a ranking has to be rather rough and ready owing to the qualitative differences between simple systems (allegro etc.) and complex systems (rallentando etc.), and the differences between features expounded on single syllables (clipped, drawled) and those expounded over polysyllabic stretches (the remainder). For all details of definition and description of features cf. Crystal (1969).

This analysis of the system of tempo is based on comparisons with several other recordings of the same speakers speaking under differing conditions on different occasions. This is because the 'tempo norm' of one speaker can be different from that of another. Five comments are relevant:

First, only 2/7 modellers (W,F) have a second version of their model which is, in one of the many possible senses, faster than their first version. The remaining modellers had slower second versions than first ones. We can attribute this effect to either (a) increasing relaxation/habituation, or (b) the effect of the (instruction to) pause for thought, or (c) both.

Secondly, the ranking of the modellers (whether we take first or second versions) is not related to the localisation of the segmental phonology of the modellers. [N]

[N] Apart from H1, who 'should' be closer to the bottom of the figure, the first versions rank the modellers better in this respect - segmental phonological localisation - than the second versions. This also, presumably may be a product of the instructions requiring pause for thought.

Thirdly, the typography of slowly on the slips appears to have elicited rallentando or drawl in 6/15 models. Four of these 6/15 are second versions, that is they have probably been elicited by the opportunity to examine the layout. (We might want to add W1 to the 6/15 on the grounds that accelerando **before** slowly may be equivalent to drawl, or lento, or rallentando **on** slowly.

Fourthly, the clipped nature of Oh in 4 /15 of the models may be attributable to the combination of its position at a line-end and its association with an exclamation mark.

Fifthly, the boundaries of 15 of the stretches of marked tempo are associated with TU boundaries (5 /15 after came, 4 /15 after said, 3 /15 after Oh, 2 /15 after later, 1 /15 after you).

There are some minority agreements amongst modellers as to the locations of features from other prosodic and paralinguistic systems as follows (cf. Crystal & Quirk 1964; Crystal 1969 for systems and features):

(a) system of rhythmicality:

of the four occurrences of glissando, three fell within the stretch a few . . . fellow came;

(b) system of prominence:

of the six occurrences of piano/pianissimo, five fell within the stretch a few . . . fellow came;

(c) system of tension:

of the four occurrences of tense, three incorporated you; all of the three occurrences of precise fell within the stretch Oh! it . . .wanted;

(d) system of voice quality:

of the ten occurrences of creak, three occurred on slowly, two on wanted, two on came.

Whether or not there are specific causes for these slight regularities, I do not know. [N]

[N] If there are specific causes, their precise nature will be even less certain. One **might** speculate that the first two categories of minor agreement (rhythmicality and prominence) - affecting as they do the first part of the sentence - were associated with the tendency of people not wanting to be overheard in language laboratories. The third category (tension) might be a consequence of the quotation marks on the slip. The fourth (voice quality) possibly has to do with falling tone, or drop plus rising tone, coming finally in the respective TUs (come, wanted) or with some dimunition of speed/prominence (slowly).

I have tried to give a picture of the variability amongst my models (and between them and QC), in order to show what range of phenomena were faced by my second group of informants when they were asked to repeat these model realisations.

Imitations, imitation variation.

A single tape of all the model realisations was compiled. The first and second models by any particular modeller occurred in sequence in the original order (first, second) on this tape. [N]

[N] The order of models , as presented to all imitators in the work discussed here, was 1 (Fa), 2 (Fb), 3 (Kna), 4 (Knb), 5 (Pa), 6 (Pb), 7 (Ha), 8 (Hb), 9 (Wa), 10 (Wb), 11 (Sua), 12 (Sub), 13 (Hua), 14 (Hub), 15 (Saa); [(Fa),(Fb) are the same as F1,F2 of fig. B54, etc.]
There do not, in fact, appear to be any obvious effects due to the order of models. (Cf. discussion of Fig. B66 below.)

The same composite tape was presented to all imitators, in other words all imitators heard the models in the same order.

It will be recalled that Quirk and Crystal presented one model sentence to 46 imitators, asking that the two halves of the model - with the break after came - be repeated. I now discuss the outcome of having asked 13 imitators to repeat 15 model realisations. The models were presented to the imitators in three forms, as follows:

(a) as a whole sentence (no break),

(b) in two parts (break after came),

(c) in three parts (breaks after came and said).

Method (a) was abandoned very rapidly, since imitators were unable to reproduce the syntactic string with any accuracy. (Interestingly, the most widespread changes were the omission of walked up slowly and, and the omission of that.)

The distribution of conditions of the models for the imitators is given in Figure B58. Note that three of the imitators (3, 4, 6) were also modellers (of models 15; 11 & 12; 13 & 14, respectively). Two of the models (1 & 2) were not presented to five of the imitators (8, 9, 10, 11, 12) because of an error on my part.

The hearing of the models, and the recording of the imitations of them took place under the same laboratory conditions with which informants had by then become very familiar. The following instructions (which were repeated) were given to imitators:

> 'I am going to play you a tape on which different people are saying the same sentence in their own way. I shall play each speaker's sentence to you (Ø)/(in --- parts) and I would like you to repeat it/them as closely as possible.'

Notice that although I am <u>calling</u> this group of informants 'imitators', to contrast them with the modellers, the instructions did <u>not</u> ask them to imitate. There are problems with whichever agentive one uses: 'repeaters' suffers different but equal disadvantages as 'imitators'. 'Respeakers' is perhaps most neutral, for the desired sense, but it is clumsy. I shall use 'imitators', but it should be borne in mind throughout that they were not asked to imitate.

Transcriptions of each model, and all its imitations, were made on separate sheets. Figure B59 gives a fragment of one such sheet. Some remarks about the transcriptions are necessary since they differ somewhat from those in Crystal & Quirk (1964), Crystal & Davy (1969) and Crystal (1975), in respect of

(a) content,

(b) layout and typography, and

(c) principle of analysis.

(a) Differences in transcriptional content.

In addition to the analysis of prosodic and paralinguistic features, the transcriptions note those places where an imitator has made use of a phonetic realisation which is judged to be

(i) non-normal for him/her, and

(ii) approximating some phonetic realisation in a corresponding part of the model sentence.

Thus, imitator 5 has a realisation [e] in the first syllable of later, where I guess that her own realisation would 'normally' be [eɩ]. In the transcription this is symbolised as /e/(eɩ). The model has [eː]. Imitator 5 also has /ɷ/(ʌ) in up, and /ɔ/(əɷ) in slowly. Missing elements in the imitations (with respect to the model) are symbolised by Ø, and have to be interpreted from the context of their location. Thus Imitator 2 omits A, imitator 4 has neither pause nor onset between said and Oh.

(b) Layout and typography.

Instead of using marginal labels for prosodic and paralinguistic effects, they are marked on a second line of transcription - mainly so that vertical comparisons can be made. The beginning of some prosodic or paralinguistic feature is marked with a single left quote mark, the end with a single right quote mark. Below the beginning of an effect its label appears in quotes, below the end, in square brackets. The only other change of notation is that onsets are overtly distinguished as to whether their pitch-prominence has the characteristics of a booster, a drop, or a continuance. [N]

[N] Criteria for the recognition of onsets are given by Crystal (1969: 143, passim). However, the three subtypes of onset proposed here, ([/,] (onset with drop), vs. [/:] (onset with booster), vs. [/.] (onset with continuance) are not allowed for in Crystal's system since it is based on the idea of a speaker's own norm. (Whereas here what is needed is a comparative frame.) Nevertheless, even in the small fragment of transcription provided (fig. B59), it is clear that the distinction is amenable to agreement (imitator 6 fellow), and disagreement (imitators 1, 2, 3, 4, and 5, on fellow).

'n.a.' is marked in the imitations at those points where either

(i) a pause in the model cannot be realised by the imitator because it occurs at a break-point in that imitator's hearing of the model;

or

(ii) a pause in the model is irrelevant because it was induced by the way in which the models were obtained (the gap between slips 2 and 3).

(c) Principle of analysis.

Sentences of the modellers are analysed in the normal way. In the case of imitators however, the simple systems of tempo and prominence (as they affect polysyllabic stretches) are analysed not with respect to the norms of the imitators' **own** prosodic realisations, but with respect to the norm established by the model. Thus, if an imitator is analysed as realising pp or allegro, it means very quiet, or fast, relative to the loudness or speed of the model.

The rest of my discussion here depends upon comparisons between transcriptions of 185 imitator realisations (15 models x 13 imitators (- 10)), and between those imitations and their models.

There is considerable variation both amongst all the imitations of some particular model, and within one imitator across different models. Here I shall only deal with variations in the disposition of TUs, variations in tonicity, variations in onset location, and variations in the tonic selection for particular tonic locations.

To begin with let us consider the variable performance of two imitators (4,9) in respect of tonic selection for the 15 models. Figure B60 gives the range and types of tonic alterations to 15 models by two imitators. The types of tone in the models (and the frequency of each) are given across the top of the table. The tones with which the imitators replace the model tones are given on the left side of the figure. The number of changes of each type is given for each imitator (4,9). The subtotals of expounded changes are added to the subtotals of model tones which are not realised at all (Ø) to give the total of changes. The totals on the right give the number of times a particular replacing tone was used by an imitator, regardless of which model tone it is replacing.

One may observe the following:

(1) Imitator 4 (Im.4) makes more expounded changes to the models than Im.9 (32/89 vs 22/89). Zero exponence ('omission' of a tone) is about equal, though it occurs for more tone types for Im.4.

(2) Im.4's favourite replacing tone (by a large margin) is RF, which replaces four tone types. Whereas Im.9's favourite replacing tone is F, replacing three tone types. The replacement of RF by F and of F by RF is symmetrical for Im.9, it is not for Im.4.

(3) Im.4 changes a significantly larger number of simple tones (R, F) than Im.9; and has a marked tendency to replace them with complex tones (RFs and FRs).

This is a somewhat gross picture, since it adds together the variation resulting from the imitations of all 15 models. Nevertheless, it characterises two important things about these two imitators:

(a) Im.4 is more variable in respect of the sum of models than Im.9;

(b) Im 4 has a marked preference for complex over simple tones compared with Im.9;

(c) Half of the changes initiated by Im.9 are of zero exponence (Ø - 'omission'), barely more than a third of Im.4's are.

In these respects at least, given the hypotheses (H1 to H4) above, I am suggesting that Im.4 and Im.9 have different underlying tonic systems, in the sense in which I used that phrase above. Credence is lent to this possibility by two observations from outside this material:

(i) the segmental phonology of Im.4 is markedly localised to Manchester, that of Im.9 is RP;

(ii) QC found for their educated speakers, almost all of whom had southern British English segmental phonologies, that the contrast between F and RF was low, and that replacements arising from this fact could be in both directions. (RF became F more frequently than F became RF, though these different distributions may have had more to do with the frequency and the syntactic and attitudinal functions of F and RF in their model than with any differential conditioning inherent in the tones. See further below.) This finding of QC's fits with observation (2) above.

A less gross picture of agreements and disagreements with model tonic choices is given in Figure B61 and in Figure B62. Figure B61 shows the distribution of agreements between imitators and the first 8 models, in terms of tonic selections at different points in the models. Thus, M1 (model 1) has a fall on later; of the 8 imitators who repeated a tone at this point, six repeated a fall (hence 6/8). The fractions in the body of Fig. B61 incorporate only those imitators who incorporated some tone for a model tone. Imitators who ignored a model tone, rather than changing it, do not appear in the denominators of these fractions. The totals at the bottom of Fig. B61 give the % agreement by all tone-realising imitators, for a specified tone of a particular tonicity (location).

Thus, for example, the overall agreement with all the models having F on later is 87%, but there is only 41% agreement with all the models having R on later.

%s appearing in round brackets are those having to do with some tone which only appears in one model.

To complete the picture of agreements given in Fig. B61, Figure B62 gives the breakdown of the %-complements from Fig. B61, i.e. the disagreements. (The minority tonicities on few, fellow, and wasn't are omitted.)

Thus, the complement of the 87% agreement with fall on later (Fig. B61, second column, below the line) is given in the topmost, leftmost cell of Fig. B62 - the disagreements are all (13%) changes to rise. There is only 41% agreement when later has a rise in the models; of the complement of this (59%), the disagreements are also varied : 26% are replaced by fall, 18% are replaced by level, and 15% are replaced by fall-rise.

Joint perusal of Figures B61 and B62 prompts several different kinds of question.

I. <u>To what extent does agreement with some particular tone in the models vary with its location in the sentence ?</u>

If we examine model tones in different places - and I restrict myself to falls and rises for this discussion - we find the following degrees of agreement amongst those imitators who repeated <u>some</u> nucleus at the points in question:

falls		rises	
fellow	100%	few	100%
came	97%	wanted	97%
later	87%	later	41%
wanted	75%	you	38%
slowly	62%	said	35%
Oh	57%	slowly	0%
wasn't	33%		

Remembering that we are excluding imitators with Ø realisations, the average agreement with all model rises is 52% and with all model falls is 73%. As can be seen, however, these averages obscure important degrees of correspondence between the rate of agreements on rises and the rate of agreements on falls. Commenting on similar kinds of averages for tonic choice, and noting the drop in overall level of agreement from that found for onset location (cf. Fig.8.63, and commentary below), Quirk and Crystal write:

> probably most speakers would agree that they are intuitively aware of less randomness in nucleus selection than [these figures] would imply. Our uneasiness is confirmed when we examine the grossly uneven way in which agreement is registered in the ... material (1966: 362).

In the light of questions II and IV below (and their answers), the word 'randomness' seems not to be justifiable. It would take a great deal more analysis of the kind represented here and in Quirk & Crystal (1966), to show that levels of agreement and disagreement in respect of tonic selection had no structure. The mistake arises from considering averages of tonic selection agreement **unconstrained by** their tonicities, and is compounded by considering variations of agreement according to tonicities ('the grossly uneven way in which agreement is registered' 1966: 362). Halliday (1967) gives categorisations in considerable qualitative detail which show why the overall (statistical) distribution of some particular tone should be thought of as comprising a set of distinct sub-distributions which may **not** be predictable from it. These distinct contributory distributions of a tone depend upon the tonality and tonicity of the parent TU, and upon the syntax, focus, key, information structure and illocutionary force of the environing clause.

From this we can argue that, given our model sentence, the willingness of imitators to repeat a given tone depends not only on the salience of that tone in respect of the underlying prosodic systems of the imitators. It depends also upon the perceived feasibility of realising some particular cluster of clausal attributes (key, focus, force, information, etc.) with that particular tone. Under such an interpretation as this, zero realisation may then have resulted from a perceived incoherence of that cluster of clausal attributes having tonicity at all. On the other hand, imitators who retain the tonicity of the model but alter the selected tone can be thought of as:

either (i) representing the same cluster of clausal attributes as the model, but expounding them through a different underlying tonic system;

or (ii) representing a different cluster of clausal attributes, or the same attributes but differently weighted, from that of the model, and therefore, on the assumption of a similar underlying tonic system to that of the model, expounding it by a different tone. [N]

[N] It is impossible, with the present material, to distinguish these two possible types of imitator. One requires, for such a purpose, to use a model sentence with different possible clusters of clausal attributes from those of this one.

There appear to be certain cases of tonicity (e.g. came in the present material) which are mandatory, but whose cluster of clausal attributes is minimal in terms of marked information value. The high agreement for cases such as these (97% amongst 13 imitators of these 8 models) suggests that the differences between underlying tonic systems expresses itself least in mandatory tonicities having very low marking function (cf. II below). If this is the case it will be important to distinguish tonic distributions according to their tonicities (B,C44).

II. To what extent are the replacing tones for a given model tone different in kind or proportions in different locations ?

A close examination of Figure B62 shows the importance of tonicity for the variable selection of a given tone. For instance, compare the replacements for F on later (13% R), and wanted (25% R) on the one hand, with the replacements for F on came (1% L, 2% RF), and Oh (4% L, 39% RF) on the other. For the different imitators represented by these figures, it is clearly the case that association between F and the clusters of clausal attributes at later and wanted facilitate one kind of replacement (R), whereas the association between F and those at came and Oh facilitate another (L/RF). Needless to say, these facilitations differ as to numerical value from one intonational variety to another.

The proportions of the same replacement tones in different locations are quite variable. This makes one view with some caution – unless it is to be constrained by such considerations of tonicity as have been suggested above – one of the conclusions of Quirk and Crystal 'it is very unusual to find it [sc. a fall] replaced by a rise' (1966: 365-6). This holds true for came, Oh and slowly, but not for later and wanted. The pattern of replacements for F on slowly seems to be very different from the patterns on later, came, Oh and wanted. Such a varied representation of complex tones may indicate a highly marked function, and further evidence for this – though not indicative of exactly what is being marked – is provided by the replacement patterns for other model tones (R, FR) in this location.

The numerical level of replacements are clearly affected by tonicity. I have discussed the variable equivalence between F and R above. The equivalence between F and RF (Quirk and Crystal 1966:

363) is also variable. 39% (of the 43%) of replaced Fs on Oh are replaced by RF; 2% (of the 3%) of replaced Fs on came are replaced by RF. 20% (of the 20%) of replaced RFs on Oh are replaced by F; but only 13% (of the 38%) of replaced RFs on you are replaced by F, the remainder being replaced by R.

These kinds of facts indicate that the repeatability of a tone by an imitator, setting aside for the moment the effect of his or her different underlying system, is a function of its position in the sentence and of the markedness potential of the bundle of clausal attributes at that point.

III Given some particular tone in some location in more than one model, how variable is the agreement by all imitators with the different models ?

If we consider all the model Fs on came (regardless of whether they are narrow or not, or whether they co-occur with a booster or not) we see that there is very little variation of agreement by imitators. Agreement ranges from 100% (M1, M2, M3, M4, M8) to

92% (M7).[N]

[N] The agreement figures in the body of Fig. B61 express agreement of tonic selections regardless of their co-occurrent range or pitch-range characteristics.

However, if we turn from Fs on came to Fs on slowly, a very different pattern emerges. Here, for model Fs (M1, M2, M4, M6), agreement ranges from 83% to 46% - respectively 83%, 62%, 69% and 46%.

There is nothing in the immediate environment of the F to account for the difference between agreements with M4 and agreements with M6. Both Fs co-occur with boosters, and both are the first elements in a FpR compound, agreement with the Rs of which in the two models only differs by 6%. M4 starts rallentando on slowly, M6 has drawl. As to the more distant environment, later carries a FR in M6, with which there is 46% agreement, and a R in M4, with which there is 69% agreement.This looks enticing as an explanation because of the parallelism of agreement levels, but it can hardly be invoked because it is in the first half of the imitation task, before the break. It seems possible that the imitators carried the effect of M6's upward glissando (which occurred in A few minutes . . . came) over onto this first tone and that this may account for the large proportion of complex tones amongst the replacements, but this can be no more than guesswork. (M6's prosodic systems are markedly Welsh.)

We can see then that the same tones in the same locations in realisations of the same sentence by different modellers can attract either

(i) something approaching **categorical** agreement (fall on came),

or

(ii) a more or less wide **scale** of agreement (fall on slowly).

These different types of agreement, I have suggested, have to do with the marking function of intonational systems.

When the marking function is obligatory, but minimally informative, agreement, even across variant intonational systems, is high (categorical).

When the marking function is non-obligatory and maximally informative, agreement is low (scalar), since these are the cases where variant intonational systems will tend to express themselves most strongly.

Scalar agreement ((ii) above) is of a kind **some** of whose variant levels can be explained by the prosodic contexts within the model realisation, but others of whose variant levels **must** have to do with differences between the prosodic systems of the model and those of the imitators.

The same kinds of arguments can be advanced from other tones. Consider for instance, the almost categorical agreement between imitators and M1, M3, M4, M5, M6 in respect of R on wanted. (Note that none of these are R of FpR, as QC's R nucleus **is** at this point.)

The levels of agreement are all 100%, except M3 with whose tone there is 84% agreement. But then, if we examine the fate of model Rs on later, what emerges is a similar pattern to that of Fs on slowly. The level of agreement with M3's R on later is surprisingly low, given the levels of agreement with Rs on later in imitations of M4 and M5.

M3 is, in addition, the only model with which there is not 100% agreement with the R on wanted. In both cases M3 uses a R co-occurring with a drop and with narrow range. However, this is insufficient to explain the low level of agreement with M3 since M5 has these modifications of later and nevertheless is accorded 46% agreement. It seems likely that the drop from 100% agreement with other models to 84% agreement with M3's R on wanted is caused by the subordination of that tone to the wide R which M3 has on you, which also suffers from low agreement (38%). No such explanation can be adduced for the low agreement with M3's R on later. I can only speculate on the basis of the rankings in Fig854 that the cooccurrence of drop with a narrow R is, in this position, particularly 'un-Northern', and that imitators, whether Northern or not, perceived an incompatibility between, on the one hand the segmental phonology and the prosodic and paralinguistic characteristics of A few . . .came, and on the other, the co-occurrence of drop and narrow with R on later.

IV. To what extent is the model tone which enjoys the highest frequency of agreement in some location also the most frequently **replacing** tone for models having **other** tones in that location ?

F has this double role at later and slowly ; L at said; RF at oh and you (in the latter case with some competition from F); R at wanted (with competition from F & RF).

What do these facts mean ?

Bearing in mind that these facts are summaries of variable imitator behaviour in respect of a sum of various models, I think we can say that for these hearer-speakers and reader-speakers of English the realisation of any given tone is dependent upon an interpretation of the clausal attributes - in the widest sense as outlined above - which operate at potential points of tonicity. The kinds of regularities which can answer question IV show that these dependencies have an overall structure which exists alongside the variant intonational systems of individuals.

To be sure, this structure is not one of discrete components but one of continuities and gradiences, but such characteristics do not diminish its structuredness, they only make it more difficult to reveal and to talk about. The overall structure of continuities which we are talking about corresponds to an abstraction which we may think of as '<u>the</u> intonational system of British English'. It must be thought of as being so abstract as to be underlying all of the individual intonational systems of actual speakers.

Having established the importance of tonicity as a framework for considering degrees of agreement and disagreement between imitators and models, I want for the moment to set it aside and to consider the simple overall distributions for the three simple and two complex tones (F, R, L, RF, FR).

Figure B63 presents the distribution of agreement with each tone across all models. The degrees of agreement with each tone are ranked from least agreement (left) to most agreement (right). It will be apparent that the agreement levels for each tone rank the imitators differently. The different rankings of the imitators are indicated in the traces and also, for ease of reading, in tabular form ('Im-rankings').

At the grossest level of observation, we see that the distributions of agreement for all of the tones have a roughly linear relationship. This is especially true if we ignore decreasing levels of agreement from the mean for each tone. In addition, though there are crossovers, to which I return, one finds similar ranges of difference between the highest agreements (on the most-favoured (RF) and on the least-favoured (R) tones) and the lowest agreements (RF, FR). The ranges of difference are 24% for the highest agreements and 33% for the lowest agreements. (If we ignored the values for FR provided by Im.10 and Im.1, then the range of the lowest-agreement end of the traces would be 28%.)

Beneath this rough linearity however there are changes of ranking of tones at differing levels of agreement ('crossovers'). [N]

[N] It must be borne in mind that we can infer nothing directly about the imitators from these crossovers, since in each case the rankings of them are different. (But cf. B65 A&B, below.) One can only get information about tones from these crossovers.

The most dramatic and important of these is that caused by the huge range of agreement with FR by different imitators. Im.10 and Im.1 seem to perceive and/or produce FR only with some difficulty. Yet FR has the third highest mean agreement level of all tones for all imitators. This indicates that the frequency of occurrence of FR will have a high diagnostic capacity in respect of different varieties of intonational system.

I have mentioned both the mean agreement and the range of agreement of FR. It is instructive to compare the ranges and means of all five tones, for which see Figure B64. I think it is reasonable to argue that a high mean agreement implies that, other things being equal amongst imitators, the tone in question is more easily repeated. [N]

[N] Since we are dealing with means, of course, other things rarely will be equal.

Put another way, such a tone is more productively salient than one with a lower mean agreement. This may or may not imply that a tone with a high mean agreement is more receptively salient than one with a lower mean agreement, since the relation between production and reception is very unlikely to be algorithmic (Parker-Rhodes 1978), and we have no way - independently of using the output of production - of uncovering the immediate products of reception.

Clearly, this is a dimension of prosodic realisation which cuts across any simple-minded notion of **discrete** variant intonational systems.

Range of agreement reflects the stability of a given tone across the set of variant intonational systems represented in these imitators. Tones for which the range is high can be expected to be diagnostic of types of variety of intonational system. Thus, for example, having a high or a low agreement rate with FR tones will be more information bearing (as an indicator of the type of underlying intonational system) than will a high or a low agreement rate with R tones. Before bidding farewell to Fig. B63 it is worth making one further observation, which arises in connection with a remark of Quirk and Crystal, viz.:

> polarity is most extreme between fall and rise: the distinction between these two has the clearest phonological status with a contrastiveness most resembling that between, let us say, voiced and voiceless consonants (1966: 365).

Though I am here ignoring considerations of tonicity - as indeed is the remark of Quirk & Crystal - I would expect the substance of their remark to be reflected in some way in Fig. B63. Even though Fig. B63 is based on imitations of fifteen models by thirteen imitators, rather than on imitations of one model by forty six imitators, one would expect the polarity to obtrude either in the form of high mean agreements for fall and for rise, or in the form of low ranges of agreement for fall and for rise, or both. None of these four expectations is at all obviously satisfied.

Are we to assume that polarity is not determinable from salience and/or from range of agreement ? (But then, how else **can** polarity be reflected in the data ?)

Or are we dealing with a group of models and of imitators whose systems are so totally different from those of the informants of Quirk & Crystal that their systems do not partake of the primary polarity of the QC tonic system ? (But were this the case, mightn't we doubt the likelihood that these two groups were speaking the same language ?)

The answer, I think, is more complicated than either of these primary questions allows. The polarity of any pair of tones must be related in some way to the two measures I have been discussing (**means** and **ranges** of agreement with model tones) (B,C45). But working out the best way to represent this relationship is complicated by two characteristics of the material here:

(i) the intonational systems of many of the imitators of this material clearly are rather different from those of many of QC's imitators (see below Figure B65 A&B);

(ii) though my discussion has been careful to distinguish those Figures in which the sum of models is being dealt with from those Figures in which patterns for single models are being dealt with, there is no doubt that the different ways of producing the models, as between QC's material and mine, and the difference between handling one model and many, are complicating factors for the immediate resolution of this problem.

So far I have been discussing what the distribution of agreements tells us about tones. I turn now to what the information in Fig. B63 can tell us about imitators. From the Im-rankings in Fig. B63 we can determine the similarity between imitators in terms of their mutually high and/or low agreement levels with different model tones. For instance, Im.10 and Im.3 each have high agreement with L tones, and low agreement with F and FR tones. They each have lowish agreements with R tones and middle to very high agreement with RF tones.
Figures B65 A&B present some of the more straightforward of these kinds of similarity in slightly more readable form. These depictions are qualitative simplifications of Fig. B63.

Figure B65A shows the relationships between the four imitators who have the lowest levels of agreement for each of the five tones. There are thus four imitators in each 'tone circle'. Some imitators appear in more than one tone circle because they have low agreement rates with more than one tone; thus, Im.1 has a low agreement rate with both model RF and model FR. Imitators who are similar in their patterns of low agreement are embraced by a dotted line. Thus Im.5, Im.7, and Im.4 all have low rates of agreement with model RF and with model R. Im.10 and Im.4 both have low rates of agreement with model R and with model F. Im.10 and Im.3 both have low rates of agreement with model F and with model FR. Thus we find that we have three groups of similarly disagreeing imitators:

(a) Im.5, Im.7, Im.4;

(b) Im.4, Im.10;

(c) Im.10, Im.3.

As is indicated in the figure, these groups overlap, representing their speakers as having **different** but not **discrete** tonic systems. In terms of segmental phonology these imitators may be labelled as follows:

Im.5 Southern British English (minimally localised);

Im.7 R.P. (non-localised);

Im.4 Lancashire (slightly localised);

Im.10 Yorkshire (slightly localised);

Im.3 Co. Durham (slightly localised).

A certain degree of match is clear, though there is absolutely no reason why variant prosodic systems should be thought of as having to be in a correlative, or in an implicational, relationship with variant segmental phonologies. As we have already seen in the work of Bazell, the postulate of linguistic solidarity remains to be proved (Bazell 1949, 1966, 1977a, b; and above Chs. 2 & 3).

Figure B65 B was constructed according to the same principles, but based on the four imitators with the highest agreement levels with each tone. According to the method of picking groups, we again find three groups: two overlapping and one simple:

(d) Im.12, Im.13;

(e) Im.13, Im.8;

(f) Im.2, Im.9.

Segmental phonologies for these imitators may be labelled as follows:

Im.12 Lancashire (fairly localised - i.e. more so than Im.4);

Im.13 Tyneside (fairly localised) [N];

Im.8 Southern British English (minimally localised);

Im.2 Home Counties (minimally localised);

Im.9 R.P. (non-localised).

[N] Im.13 is the only imitator who is younger than the rest (about ten years younger). I am very grateful to Ms. Emma Pellowe for undertaking such a tedious task so seriously.

Given both the simple method of forming the groups, and the fact that the distributions from which they were drawn (Fig. B63) ignore tonicities which have been shown to be important, the 'appropriacy' of these groupings, though encouraging, may be misleading!

So far, discussion has been based either upon the variable performances of the imitators in respect of all the models taken as a lump, or upon differences between the models taking the performances of all the imitators as a lump. However, the hypotheses (H1 to H4) which I advanced above, and several of my remarks about preceding Figures imply that the performance of individual imitators will differ in different models. Such differences are displayed in Figure B66 which gives a ranking of imitators against a ranking of models, showing an index of tonic alteration.

The imitators are ranked along the top of the table according to the overall consistency with which they realise model tones in a different way from the model. Most different are on the left, least different are on the right. The imitator ranking index is arrived at, for any given imitator, by taking the number of models for which the overall index of tonic alteration by that imitator is greater than or equal to 0.50, and dividing by the number of models available to that imitator.

The models are ranked down the left of the table according to the overall consistency with which they elicit changes from the imitators. The model ranking index, for any given model, is calculated by taking the number of imitators whose index of tonic alteration is greater than or equal to 0.50, and dividing by the number of imitators of that model.

High values of the imitator ranking index (IRI) indicate that the imitator is consistently different from the sum of the models in tonic realisation. Low values of the IRI indicate that the imitator is consistently similar to the sum of the models in tonic realisation. High values of the model ranking index indicate that a model has tonic realisations which are less amenable to agreement by all imitators than a model with a lower model ranking index. The labels 'harder', 'easier', (of models) and 'further', 'closer' (of imitators) are used for ease of reference: they do **not** have theoretical or evaluative connotations.

The body of the table contains an index of tonic alteration between every pair (i,j) of all imitators and all models. The index of tonic alteration is calculated by dividing the number of changed model tones by the total number of model tones. Within what was counted as a change differential allowance was made for:

(a) zero realisation by the imitator;

(b) insertion of tones by the imitator at points of zero tonicity in the model;

(c) imitator realisations of e.g. F of FpR for a plain F in the model.

All three of these indexes - imitator ranking index, model ranking index and index of tonic alteration - have a potential range of 0 to 1. Values of the index of tonic alteration which are ≥ 0.50 have been shaded. Cells in the body of the table having a diagonal bar are those in which the imitators are imitating their own models (self-imitation).

The numbering of the models (in the column to the right of the model ranking index) is the order of presentation of these models to all imitators. There is clearly no ordering effect in relation to 'harder'/'easier' models. That is, imitators do not increase their degrees of agreement as they hear more and more models. However, there _is_ evidence that a modeller's second version (_b_ in each case) is more amenable to higher levels of agreement than that modeller's first version (a in each case). (The only reversal of this is where the model ranking index for Hut b is higher than that for Hut a.)

Though we might wish to attribute this to a 'tuning' effect, we have to bear in mind that the first and second versions of some modellers are quite different (cf. Fig. B54 above). Given such differences between a modeller's a and b versions, and given the overall raising of agreement scores for the b version, it is important to gain rather more precise information about the processes behind such tuning effects, since the effects themselves are doubtless differentially distributed across speakers with different underlying prosodic systems. There is one other possible explanation of this a/b pattern which arises from the facts presented in Fig. B57; and that is that all modellers (apart from Wil b and Full b) had second versions which were, in some sense of the word, **slower** than their first versions. (This does not account for the Hut a/Hut b problem. But we have noted elsewhere that Hut's models are odd.)

In the unrecoverably surreal jargon of our days we may wish to see Fig. B66 as 'a squish' (cf. Ross (1975) on a squish which doesn't seem particularly 'squishy' to me). Alternatively, we can interpret it within the prior, and inferentially richer, framework of Quirk's (1965) notion of serial relationship. Or we can view it as an example of a dissimilarity matrix for which many well understood methods of transposition and interpretation exist (Sokal & Sneath 1963; Sneath & Sokal 1973).

If we do take Fig. B66 to be a dissimilarity matrix, we must acknowledge that it is one of an unusual kind insofar as the sets of 'objects' along the two sides of the matrix are not the same. On the assumption that the index of tonic alteration is a reflection of the difference between the underlying tonic system of the imitator and the underlying tonic system of the modeller, imitators are similar to each other insofar as they share similar patterns of values - in respect of the different models - for the index of tonic alteration (B,C46).

On this criterion Im.2 and Im.11 are very similar. On the grounds of a different <u>pattern</u> of similarities, Im.6 and Im.7 are also similar. Fig. B66 manifests variant prosodic systems as grouped, but grouped in terms of greater or lesser degrees of a **continuous** measure of similarity. To be sure, the similarities are of the same <u>kind</u> as those in Figs. B65 A&B, and, like Figs. B65 A&B, they do not take account of varying tonicities. Unlike Fig. B65 A&B however, these similarities are based on relationships with <u>single</u> models rather than the <u>sum</u> of models. Thus for example the similarity between Im.11 and Im.2 in Fig. B66 is not reflected in Fig. B65 A&B. This does not mean that one or the other type of similarity is spurious. On the contrary, it means that the relationships between variant prosodic systems are hierarchic and that neither level of that hierarchy is one-one predictive of the other (cf. Pellowe & Jones 1978, 1979; Appx. D below & Chs. 2 & 3)(B,C47). This certainly fits with graph theory and nearly decomposable systems (Harary 1969; Simon 1970; Rijsbergen 1971).

The non-implicational nature of different patterns of similarity across different underlying prosodic systems is further demonstrated by Figure B67 which depicts a comparison of values for the imitator ranking index (cf. fig. B66) against an index of range of exponence of model nuclei and against an index of overall deviation from models.

Imitators are ranked left to right for decreasing values of the imitator ranking index (IRI), against which are plotted their values for the exponence range index (ERI) and the overall deviation index (ODI). [N]

[N] Recall that the <u>lower</u> an IRI value, the <u>closer</u> that imitator is to the sum of the models.

The ODI for any given imitator is simply derived from the mean of all of the tonic alteration indices (Fig. B66) of that imitator. The ERI for any imitator is an overall measurement of that imitator's proclivity to replace tones with a <u>variety</u> of tones. It is calculated by adding the number of tone types with which that imitator replaces model F, to the number replacing R, to the number replacing L, to the number replacing RF, to the number replacing FR, and dividing this sum by 50. [N]

[N] Alternative bases for an ERI were tried:
(a) including all tones in the typology, and
(b) including all model tones and tones inserted by the imitators where models had Ø tonicity.
Both alternatives showed the same distribution, though at a lower level, as the ERI which is being defined on this numerator and is used in Fig. B67.

The comparative distributions of the IRI, the ODI and the ERI give a clear overall indication of the extent of differential similarities arising from the underlying prosodic systems of these imitators.

Two general points are worth emphasising.

First, the distributions of the IRI and of the ODI show quite strong agreement down to a value of about 0.33. At that point, for sharply declining values of the Imitator Ranking Index, the Overall Deviation Index continues to be more or less linear.

Although an imitator can occasionally realise **total** agreement of nuclear exponence with a single model (i.e. an index of tonic alteration of 0.0 in Fig. B66 - cf. Im.13 on M8, or Im.9 on M4 in Fig. B66), the overall capacity of imitators to realise a range of models does not improve below a value of 0.32 for the ODI. We have already established that there are no ordering effects on the index of tonic alteration, and therefore, though it would be perfectly comprehensible, the divergence between the IRI and the ODI cannot be due to exhaustion or any other extrinsic variable. The divergent linearity of the ODI (in respect of the three 'closest' imitators) must therefore be due to:

(a) some factor in the range or in the difficulty of the models (but is this universal, or specific to these models ?);

(b) some factor inherent in the competence of the imitators (but is this universal (the business of imitating) or specific (these particular undergraduates) ?);

(c) some interaction between (a) and (b). [N]

[N] These are particularly thorny questions and I shall not go further into them here. It is not simply that further pondering of Fig. B67 might produce more speculation - which it no doubt could - but that one needs to develop a limiting interpretive framework for these problems. This is started on in Pellowe (in preparation, a).

Secondly, the differences between the ODI and the ERI are very informative about the differential kinds of similarity in variant prosodic systems. Thus if we compare Im.5 with Im.1 we find that they both have a high overall deviation from the sum of the models (0.42), but that whilst Im.5 has the highest ERI of all imitators (0.36), Im.1 has the lowest (0.20). That is, whilst Im.5 deviated from the models by using a wide range of replacements for each model tone type, Im.1 did so by using very few tone types to replace the whole range of model tone types. We can expect the underlying prosodic systems of imitators having **wide** differences between exponence range and overall deviation to be rather distinct from those having **narrow** differences. For example, we would expect Im.1, Im.10 and Im.7 on the one hand to have rather different underlying prosodic systems from Im.3 and Im.12 on the other.

Much of what I have been discussing so far concerns different types of relationship which obtain between different patterns of preferential realisation of tones (variant underlying tonic systems). But now, briefly, I want to return to some comparisons of the present material with that of Quirk and Crystal (1966).

In discussing overall levels of agreement with various features of their model, Quirk & Crystal found that the entity attracting the most agreement was the tone unit itself. That is, regardless of where imitators choose to locate onsets, or the nucleus itself, or what the pitch movement of the nucleus is, what they agree most upon are the stretches of utterance which are to carry these entities. The second highest level of agreement was the location of nuclei (tonicity), regardless of their pitch movement. The third highest level of agreement was the location of onsets. The fourth highest level of agreement, whose relatively low level they were disturbed by, is the actual pitch movement of the nucleus. When the numbers for the exponence of nuclei were adjusted (e.g. counting high booster+wide F as a variant of RF) the level of agreement reached what was considered to be an adequate level (Quirk & Crystal 1966: 362-3). [N]

[N] The problem with this approach is that there is no external criterion by which one can determine whether one has done sufficient adjusting of the nuclear exponence. Figure B68 plots the levels of agreement before and after adjustment in the Quirk & Crystal treatment. Clearly, if the motive for the adjustment is predominantly numerical, then what has been found as the solution is the reduction of a dogleg/curve distribution of agreements to a linear one. But do we have any **reason** for supposing that such types of agreement must bear a linear relationship to each other ? And what could such a reason look like ? (See also discussion of Figs. B69 & B70 below.)

Figure B69 compares the levels of agreement for tone units, tonicities, onsets, and true and adjusted nuclear exponences in the Quirk and Crystal material and in this material. (I only consider the first five of my models.) % agreements for tone unit (TU), tonicity (TC), onset (ON), true exponence (EX^t), and adjusted exponence (EX^a) are given. Except for the fourth (EX^t), the % agreement of each of the entities is ranked with respect to the others. The combinations of features which have been used to convert EX^t to EX^a are given in note form following Fig. B69.

These combinations of features vary from model to model, but most of them accord with regularities in the findings of Quirk & Crystal (1966). N

N The motivation and the rationale for their adjustment of the true exponence scores (Quirk & Crystal 1966: 362-3, passim) are well founded, given the aims of their experiment. However, conflation of such variants defeats one of the main aims of this thesis. It is likely that the distribution of such variants is structured according to the different underlying systems which are their source, and consequently that QC's 'uneasiness' about the (low) levels of EX^t scores is misplaced.

Figure B69 is presented in pictorial form in Figure B70.

These figures raise two points. First, in the light of my five models, the first ranking - of agreements on tone units - in QC seems atypical. This is shown by the fact that the highest ranked entity in four out of five of my models was tonicity (location of nucleus). I suspect that rather than the QC level of agreement for TU being too high, their levels of agreement for TC are too low. It is possible that this can be explained in terms of the unusualness of the realisation of the QC model (see N above p. 322), but this can be no more than speculative.

Secondly, whatever the cause, or causes, for the differences between levels of agreement in QC and in these five models, it is clear that there is considerable variation in levels of agreement with these entities amongst these five models. This can be interpreted as further evidence that the similarity between variant prosodic systems is variably continuous and is hierarchically differentiated.

One possible line of interpretation is certainly blocked; it might be thought that since M3 in Fig. B70 is most similar in patterns of levels of agreement with the QC model, that that modeller has a non-localised variety similar to that of the QC model. But M3 (Kna a) had lived in Ryhope, Co. Durham all of her life, the daughter of Co. Durham parents whose own varieties she characterised as fairly broad Co. Durham - slightly broader than her own.

All of the above interpretations can be taken as evidence in favour of a general view that the similarity between variant prosodic systems is variably continuous, is hierarchically differentiated, and is not predictable from the co-occurrent phonemic/phonetic characteristics of the variety.

<u>Summary of attempt to establish variant intonational systems experimentally.</u>

The models showed considerable variation. Nevertheless, my method of obtaining the models retained some important bases for comparison between them and that of QC. Similarities between my models and that of QC were mainly in respect of tonality and tonicity. Variation amongst my models in terms of additional tonicities (compared to QC), tonic selection for given items, and types of change between first and second versions which were provided by modellers showed relationships with a crude segmental phonological scale of 'more-Northern'/'less-Northern' (Fig. B54). The typicality of the QC model in terms of my models was shown to be variable when one examines tonic selection across different tonicities (Fig. B55). Tonicity, tone units, onsets and pauses varied both amongst my models and beteen the sum of my models and that of QC (Fig. B56). Tempo variation and the distribution of some paralinguistic features showed some regularity amongst the modellers (B57). Such regularities may be interpreted as arising from the manner of eliciting the models. It is important to know the range of variability amongst the models in order to interpret the performance of the imitators who are attempting to repeat them.

Imitators were presented with fifteen models of the sentence. Imitators occasionally modified their segmental phonologies in their attempts to reflect the intonational realisations of the models (Fig. B59). A simple comparison of the realisations of two imitators (one Northern, one RP) in respect of the sum of the models showed significant general differences of agreement with tonic exponence (Fig. B60).

It was demonstrated that tonicity plays an important part in the distribution of the agreements (Figs. B61 & B62) with tonic exponence and that averaging of distributions across tonicities, under certain interpretive frameworks, was misleading. Imitators having zero exponence for a given model tone, and imitators having a wide range of replacement tones for a given model tone were both kinds of evidence which are taken to indicate that variant intonational systems express themselves most significantly at points of tonicity which are non-mandatory, but which are maximally marked. Agreements seem to be either categorical or scalar in type. The former are associated with a cluster of clausal attributes almost mandatorily carrying tone, but bearing a minimum of markedness. The latter are associated with clusters of clausal attributes **not** mandatorily carrying tone, but having high markedness potential (Figs. B61 & B62). The most favoured tone in a given position in all the models is often the most frequently replacing tone used by the imitators for an atypical model tone used in that position (Figs. B61 & B62). The presence of such patterns as these, alongside those indicating the existence of variant underlying intonational systems in individuals or grouops, are reflections of the existence of more abstract, possibly singular, intonational systems of British English.

Distributions of agreement with model tones shows that there is a roughly linear relationship between the agreements for all simple and complex tones (Fig. B63). The relationship between the **polarity** and the **conditionability** of tones, as discussed by Quirk and Crystal, and the **mean** and the **range** of agreement on tones which is discussed above is not straightforward, but it seems that the rankings given by the realisations of the QC model may not be typical (Fig. B63, B64, B70).

Tones for which agreement has both a high range and a high mean across all imitators are ones which are diagnostic of different underlying intonational systems. However characteristic groupings of imitators (Figs. B63, B65A, B65B) are taken to show that variant intonational systems are not discrete from one another. Groupings on the basis of tonic exponence **did** relate to segmental phonology (Figs. B65A, B65B) but not in a 1:1 (implicational) manner.

When individual ranked imitators are related to individual ranked models in terms of an index of tonic alteration, a serial relationship of underlying tonic systems emerges which also reflects similarities amongst imitators (Fig. B66). It is clear however that the similarities are reflections of hierarchically different aspects of the variant prosodic systems, and that different levels of the hierarchy are not 1:1 predictive of one another (Figs. B66, B67, B69, B70).

F. Analyst consistency.

Just as the model of Chs. 2 & 3 predicted the likelihood that analyses of phonological material, both phonemic and phonetic, would differ according to its general principles (see above, Appx. A,6), so it predicts that analyses of prosodic and paralinguistic features will differ in accordance with the place of the participating linguists in some VSp. Clearly, the need to calibrate such differences is no less in the intonational case than it is in the segmental case, since the interpretation of the contents of the VSp (that is, the patterns of the dispersed sampled speakers) will need to be informed by such calibrations.

Nevertheless, it would be a serious misinterpretation of the approach to imagine that calibration of analysts is tantamount to a regularising of the data. What can be learned from analyst differences (**not** discrepancies in a pejorative sense) is something about the nature of the linguistic systems which are being analysed, and the processes by which their realisations, and the differences of underlying systems which produce those realisations, are normally apprehended in daily life.

Pellowe (1974a) calibrated three analysts.

Regrettably both the original analyses (D.C., J.P. & V.M.) and the raw score tables of agreements and differences are now lost. [N]

[N] I am grateful to Professors Randolph Quirk and David Crystal for having provided a copy of one of their tapes, and a copy of its analysis by D.C. They are in no way responsible for, and may not be in agreement with, my use of their material.

The tape was a recording of three educated, male speakers in an unscripted radio discussion of the social, moral, political, and economic desirability or undesirability of corporal and capital punishment. The discussion is occasionally incoherent and often heated.

Without looking at D.C.'s analysis, which had been written on a running typescript of the words on the tape, I asked a typist to copy the typescript. The tape was then analysed onto copies of this 'blank' typescript, according to the categories of tone and of pitch range as specified in Crystal & Quirk (1964), independently of each other, by both Pellowe (J.P.) and Vince McNeany (V.M.). Raw score tables of agreements and differences between all three analysts were prepared. The frequency of all the types of agreements and differences were characterised as to prosodic/paralinguistic environment on the raw score tables.

All that remains of this work is a summary of the main kinds of differences between the three analysts. (Perhaps the most remarkable thing about these differences is that they do **not** appear to fit with the kinds of stances I have adopted in section E, above. That is, my assumption that differences between imitators, or, in the present case, differences between analysts, will be reflections, at least in part, of differences in the different underlying (varietal or lectal) systems of the analysts.)

The analyst differences fell into four classes of two types: (a) simple, (b) complex.

The two classes of simple differences were as follows.

(a,i)

One analyst hears a lone pitch range feature, the other hears no such feature. The most frequent exemplar was booster vs. ø. This was a difference which occurred in both directions for both pairs of analysts. I.e.

perception:	booster	ø
	{ D.C.	J.P.
	{ J.P.	D.C.
analysts	{ V.M.	J.P.
	{ J.P.	V.M.

(a,ii)

One analyst hears a tone, the other doesn't. The most frequent exemplars of this difference were either F or L vs. ø. The differences occurred in both directions for both pairs of analysts. Note that in these cases there were **no** 'compensatory' analyses of the type that we might be led to expect from the material above on imitator behaviour (section E), or from Quirk & Crystal (1966). That is the analyst, of a given pair, who heard no tone did **not** have an analysis involving a pitch range feature at that point (see (b(i)).In other words these differences (a(ii)) are not differences in the perception of kinesis.

The two classes of complex differences had to do with kinesis and contrastivity, respectively.

(b(i))

Both analysts hear a pitch difference, but whilst one of them hears it as static (pitch range) the other hears it as dynamic (tone). The most frequent exemplar of this difference, for both analyst pairs was booster vs. R.

(b(ii))

Both analysts hear dynamic pitch change but they differ as to its precise nature. For D.C.:J.P. the commonest exemplar was R vs. FR, for V.M.:J.P. the commonest exemplar was F vs. L, a difference which occurred in both directions.

It is a pity that more detail is not available for discussion here. But clearly this little which has remained after the loss indicates that more work should be done on prosodic analyst differences.

In considering the structure of prosodic variation as uncovered by either the method used in Sections C & D above, or that used in Section E above, it will be important for us to bear in mind that any final representation of a speaker in a total computed VSp will have to take into account such analyst differences (and agreements) as are barely sketched here.

Appendix C

Establishing paralinguistic, syllabic, morphological, & syntactic criteria

The general distributional reasons for having to express some criteria as quantitatives, and the problems associated with that practice, have been rehearsed at some length in Appendix B above, and from the point of view of taximetric considerations in Appendix D below. In the list of criteria which follows, the capital letters in the left hand margin are simply counters, or labels for various kinds of raw totals.

There are various bases for defining the following criteria which are almost all percentage ratios.

Either (i) they are defined as they are for the sake of comparison with research on 'educated varieties' (Quirk et al. 1964);

or (ii) they are defined on the basis of unpublished piloting by people associated with the Tyneside Linguistic Survey;

or (iii) they are defined on the basis of the intuitions of single hearers (some of whom were non-native speakers and some of whom were native speakers of the particular varieties concerned);

or (iv) they are defined on the basis of blind guesswork.

Note that even guesswork which is said to be blind may contain a substantial component of unarticulated, or unrealised, thrust from theoretical expectations or desires. Facts neither prompt conjectures nor do they refute them. On the contrary, those things which one admits as facts are shaped by and dependent upon only and precisely one's conjectures (C,C1).

In the criteria specified below one does not, in most cases have enough, if any, information about the distributions or cooccurrences of the variants in <u>any</u> population of speakers of English. There is certainly not enough to be sure what the most appropriate arithmetical base for the various criteria should be. For example, in criteria 336 to 342, and 349 to 352, which express the frequency of occurrence or cooccurrence of various prosodic and paralinguistic features, I have used the base of "total TUs" in the denominators, and "number of TUs wholly or partly affected by [the given paralinguistic/prosodic feature]" in the numerator. Whereas in criteria 343 to 348, which express the frequency of occurrence of the same <u>types</u> of features, I have used the base of "total number of words" beneath a numerator of "number of words affected by [the given feature]".

Other things being equal, these different bases will provide rather different levels of measurement for the given features. (One would guess, for instance, that if one measured the frequency of occurrence of a given paralinguistic/prosodic feature on <u>both</u> of these bases, then one would find much higher levels of the measure in the first case than in the second.) But one must pay the price of this statistical variability in order to find a base which is overall the least biassing, whilst giving reasonable discrimination between speakers. It seems to me that if there **were** such a single best base it could not be known in advance of experimenting with several possible bases (C,C2).

OB Total no. superordinate TUs (as defined by Crystal 1969).
AA No. of high & extra high boosters which co-occur with forte or fortissimo syllables. (Cf. Crystal 1969: 300, passim)
336 (AA/Z2+Z3)100
[Z2+Z3 is the total number of high & extra high boosters derived from the matrix on p.212, Appendix B.]
BB No. of non-nuclear drawled syllables preceded by a booster.
CC No. of clipped tones preceded by drawled boosted syllables.
337 (CC/BB)100 (Cf. Crystal 1969:154)
DD No. of superordinate TUs carrying in whole or in part allegrissimo, allegro, lento, lentissimo, accelerando, rallentando.
338 (DD/OB)100
EE No. of fortissimo syllables.
FF No. of fortissimo syllables which are nuclear (Crystal 1970).
339 (FF/EE)100
GG No. of superordinate TUs carrying in whole or in part the terms piano, pianissimo.
340 (GG/OB)100
HH No. of superordinate TUs carrying in whole or in part the terms forte, fortissimo.
341 (HH/OB)100
JJ No. of superordinate TUs carrying in whole or in part the term low from the pitch range system.
342 (JJ/OB)100
KK No. of words in stretches of spiky (up), spiky (down).
343 (KK/OA)100
[OA is total no. of words.]
LL No. of words in stretches of glissando (up), glissando (down).
344 (LL/OA)100
MM No. of words in stretches of staccato.
345 (MM/OA)100
NN No. of words in stretches of legato.
346 (NN/OA)100
PP No. of words in stretches of slurred, lax.
347 (PP/OA)100

QQ No. of words in stretches of tense, precise.
348 (QQ/OA)100
RR No. of superordinate TUs carrying in whole or in part breathiness.
349 (RR/OB)100
SS No. of superordinate TUs carrying in whole or in part huskiness.
350 (SS/OB)100
TT No. of superordinate TUs carrying in whole or in part creak.
351 (TT/OB)100
UU No. of superordinate TUs carrying in whole or in part laugh, giggle.
352 (UU/OB)100

In the next group of criteria, I am concerned to distinguish speakers in terms of the extent to which they mark particular word boundaries in a quite specific way. It must be emphasised that any inter-speaker similarities accruing from not marking boundaries in this way do not necessarily imply a similarity of realisation. The possibility of this effect is not particularly harmful and, anyway, reflects in part the general theoretical approach of Chapters 2 and 3 above.

VV Number of possible slots for linking and intrusive 'r' which are not realised by a glottal stop.
WW No. which are so realised.
353 (WW/VV+WW)100
XX Total word-final /t/ = [t], V____**V
e.g. [bɪt əv], 'bit of'
ZZ Total word-final /t/ = [ʔ], V____**V
e.g. [bɪʔ əv], 'bit of'
AB Total word-final /t/ = [ɹ], V____**
e.g. [bɷɹ eɪ], 'but I'
(See also AG below ([d]--->[ɹ].)

McNeany has suggested, in conversation, that the grammatical category of the item is important here; that nouns do not usually undergo

this process, and that the items which do experience the process are usually verbs. Thus [gɛɹ ɒf], [nɪɹ ɪm wɷn], [pɷɹ ɒn jə hɛ̞t̚ʔ] 'get off', 'knit him one', 'put on your hat' are more probable than [pɷt̚ʔ jə hɛ̞ɹ ɒn], [ɪmɪteᵊt ə kɛ̞ɹ əgɛn], [pɪʔ d̤ə baɹ ɷp̚ʔ] 'put your hat on', 'imitate a cat again', 'pick the bat up', though the latter are by no means impossible. Though this point seems to me to be right, it leaves out of account both the widespread productivity of the process, and the fact that it occurs, without seeming odd or marginal, in form classes other than these two, cf. [bɷɹ eɪ], [lɷk ɪɹ ɪz], [eɪ nɛvə nɪu ɪʔ wəz wɛɹ ət̚tə eɪsɹɪŋk̚ʔ] 'but I', 'look at 'us'(=me)',
'I never knew it was wet at the ice-rink'. The phonetic constraints on the occurrence of the process are that the following item must begin with a vowel and that the /t/ or /d/ cannot be preceded by a diphthong.
Thus *[ʃɛɷɹ ɪt̚ʔ], *[leɪɹ ɪt̚ʔ] 'shout it', 'light it', are both impossible forms. The second constraint may in fact include all lax vowels, not just diphthongs, thus ?*[ʃuːɹ ɪt̚ʔ], ?*[hiːɹ ɷp̚ʔ], 'shoot it', 'heat up', are probably impossible, though in formulaic phrases long vowels may precede the /t/ or /d/, as in ?[uːɹ ə tuːn] 'out of town'.

AC Total word-final /t/ = [d], V____**V
e.g. [gɒd əp], [bɷd i], [gɒd əm], 'got up', 'but he', 'got 'em'

It might be argued that the third example, got them, was a case of regressive similitude, (cf. AX below; [wɪz dɹeɪv] for [wɪst dɹaɪv]), but this could hardly be argued for the second example: [bɷd i] (=$_{NL}$ [bʌt (j̥)i]).
My guess is that for a significant number of speakers this voicing is independent of assimilation and similitude.

354 (ZZ/XX+ZZ+AB+AC)100
356 (AB/XX+ZZ+AB+AC)100
357 (AC/XX+ZZ+AB+AC)100
AF Total word-final /d/ ≠ [ɹ], ____**{V}/{/h/}...
AG Total word-final /d/ = [ɹ], ____**{V}/{/h/}...

e.g. [wʊɹ i], [hɪɹ ɪn] 'would he', 'hid in'
(Cf. remarks concerning AB above.)

358 (AG/AF+AG)100

AH Total word-final /n/ ≠ [m], ____**{/m/}...

AJ Total word-final /n/ = [m], ____**{/m/}...
e.g. [dʊm mɔ], [bam mɔtʃəz] 'done more', 'ban marches'

AK Total word-final /n/ ≠ [ŋ][1], ____**{/k/}/{/g/}...

AL Total word-final /n/ = [ŋ], ____**{/k/}/{/g/}...
e.g. [fʊŋ kɪŋ], [dʊŋ kɛʊ], [dʒɔmɪŋ gɔl], [θɪŋ gɹup]
'fu<u>n</u> king', 'du<u>n</u> cow', 'Germa<u>n</u> girl', 'thi<u>n</u> group'.

---

[1] That is, [ŋ] ± following [ʔ].

---

359 (AJ/AJ+AH)100

360 (AL/AL+AK)100

ED Total of word-final consonants = Ø

361 (ED/total number of words)100
e.g. [ə], [hɛ], [fɪ(j)ʊp], [blɒn gɔl] 'of', 'have', 'fill up', 'blon<u>d</u> girl'
(the last may also be coded under AK, AL).
Deletion means deletion without any form of substitution (cf. AB, AC, AG, AJ, AL), except, possibly, vowel lengthening, glides or epenthetic vowels as in [fɪjʊp] for 'fill up', [howɪnɪʔ] for 'hole in it', etc.

AP Total word-final /v/ ≠ [f], ____**{/t/}/{/s/}/{/f/}...

AQ Total word-final /v/ = [f], ____**{/t/}/{/s/}/{/f/}...
e.g. [hɛf tɪ] [hɛf sʊm] [gɪf tɪ] [liːf fɔ] 'have to', 'have some', 'give to', 'leave four'

362 (AQ/AQ+AP)100

363 Number of boundary metatheses/total number of words.
This phenomenon is somewhat elusive analytically, but quite obvious perceptually. For the sake of convenience it is transcribed as a [ʔ] -
e.g. [eɪʔtin] [ðɪʔseᵊdʒ] 'eighteen', 'this age'.

AR Number of vowel ± consonant deletions within word (syllable

deletion) replaced by syllabic [n̩±ʔ]
e.g. [bɛnʔn̩], [swɪmn̩] 'Benton', 'swimming'

AS Number of vowel ± consonant deletions replaced by syllabic [m̩±ʔ]
e.g.
[bɒʔm̩]/[bɒ(ə)m̩], [stɒʔm̩/[stɒm̩], [snɔjystɒm̩m̩m̩mãn]
'bottom', 'stopping', 'it's no use stopping him man'

AT Number of vowel ± consonant deletions replaced by syllabic [l̩±ʔ]
e.g. [ɪvɛntʔʃl̩ɪ] 'eventually'

AU Number of vowel ± consonant deletions in which the syllable is not restored
e.g. [laɪbɹɪ] [gɒʔ ðə dɹɛs] 'library', 'got the address?'

364 (AR/AR+AS+AT+AU)100
365 (AS/AR+AS+AT+AU)100
366 (AT/AR+AS+AT+AU)100
367 (AU/AR+AS+AT+AU)100

AV Number of progressive assimilations with or without vowel deletion $\{ VC_1(V_1)^{**}(V_2)C_2 \dashrightarrow VC_1(Ø_1)^{**}(Ø_2)C_1 \}$

e.g. no V deletion [ät θə päʔk] 'at the park'
V_1 deletion [goɷn tθ(ə) päʔk] 'going to the park'
V_2 deletion [peɪntfɛks] 'pint of Ex(hibition Ale)'

368 (AV/(0.5 x no. of words))100

Notice here the change in the denominator. Since it is not known (by me at least) which subset(s) of lexical items are candidates for the process expressed in this criterion, and subsequent criteria, the denominator cannot be the number of such candidates. The only remaining denominator is the total number of words, since these are processes which affect words. But this number, by being so high relative to the numerator, will make all the measures very low, and hence of hardly any importance in the classificatory procedure. Half the number of words is therefore used since first this will increase the measures in a uniform way across all speakers, and secondly many of the processes reflected in these criteria are ones taking place between two words.

AX Number of regressive similitudes with or without vowel deletion.
e.g. [wɩz dɹaɩv] 'whist drive', [sãdːə] 'Saturday'
One should perhaps code the first of these examples as a double occurrence of regressive similitude, since one must account for the /t/ deletion before one can consider [d] to be having regressive effects on the similitude of [z]. Thus [wɩst dɹaɩv] --> ([wɩsd̯ dɹaɩv]) (AX_1) -->([wɩzd̯ dɹaɩv]) --> [wɩz dɹaɩv] (AX_2); however since such intermediate realisations do not occur, I do not propose to code such examples twice over.

369 (AX/(0.5 x no. of words))100

AY No. anaptyxised vowels
e.g. [ɷmbəɹɛlə] 'umbrella', [fɩləm] 'film'

370 (AY/(0.5 x no. of words))100

AZ No. of shifts of primary word-stress
e.g. [nəˈbɒdɩ] 'nobody', [ãgɹəˈveɩtn̩] 'aggravating'

371 (AZ/(0.5 x no. of words))100

FH Total miscellaneous other elisions

372 (FH/(0.5 x no. of words))100
e.g. [gɩz] for [gɩv əz] (i.e. 'give me')
It might be inferred from the example that this should be simultaneously coded as progressive assimilation with vowel deletion ([s] to [z], [ə] to Ø), i.e. AV above, plus final consonant deletion ([v] to Ø), i.e. ED above.
However, there is simply not enough information about the nature and extent of the distribution of such processes as are represented by AV & ED for one to be able to guess the likelihood of their co-occurrence, and hence for one to have an estimate of the conditions for, and the significance of, such a co-occurrence. When this kind of doubt arises, as before, my theory indicates that one should specify two criteria, and then permit the dynamic nature of the space to tell one more about their relationship.

BC Number of vowels which are nasalised of a random selection of 10 V **/______{N}, and 10 V **{N}/______;

where {N} is any nasal.

373 (BC)5

BD The number of occurrences of "perceptually significant" consonant length which does not involve assimilation or intermediate deletion (of the kind outlined in the note to AX).
E.g. [jɛsː] 'yes', [ɛk̚ʔːspɛkt] or [ɛkːspɛkt] 'expect'

374 (BD/(0.5 x no. of words))100

FA No. of epentheses of (usually homorganic) consonants after liquids and nasals.
E.g. [ɛlt̚ʔs] 'else', [tʃɒmpskɪ] 'Chomsky'

440 (FA/(0.5 x no. of words))100

FB No. of deletions of homorganic consonants after liquids and nasals.
E.g. [stãms] 'stamps', [fiːlz] 'fields'

441 (FB/(0.5 x no. of words))100

The next group of criteria explore the representation of 'non-standard' grammatical and lexico-grammatical contrasts in the population. It will by now be obvious that different categories of criteria have differential degrees of localisedness attaching to them. That is, some vocalic PDVs are very specifically localised to Tyneside, other variables are represented throughout the North-East, others north of the 'Bristol-Wash' line, yet others throughout the whole of the UK. It is extremely important to emphasise that an accountable method of examining linguistic diversity must incorporate such variability of the socio-spatial distribution of criteria if it is to have anything credible to say about the differential salience and significance of spoken variants.

BF^1 No. of clause enclitics realised non-emphatically
E.g. [eɪl 'manɐdʒ bət] 'I'll manage but'.

BF^2 No. of clause enclitics realised emphatically,
E.g. They're useless '<u>them</u>
My skirt's too short '<u>this</u>
I could just go a toasted sandwich '<u>me</u>
[Often, but not necessarily, the enclitic is tonic; if so,

then usually it is the second (often rising) part of a compound tone.]

BG No. localised verb forms (tense, modals, morphology).
E.g. 'he's went' ['he has gone'], 'I come' ['I came'], 'he had sang' ['he had sung'], 'I might could manage it' ['it's possible that I could manage it'], 'I've cooken dinners for so many' ['I've cooked . . .'], 'I'll better get off' ['I'd better get off'], 'I've senden Harry' [' . . sent'];
and heard by Vince McNeany:
(Sharron) 'You used to could buy them frozen, I mean ɛ̃: you could used to.' .
(Christine correcting S.) 'You used to could man' ['used to be able to'],
'That's what happens when you be naughty' [' . . are . .'], 'When I dance it makes my cough be worse' [' . .seem/get/?become . .']
"HAVE a chance to be +V-ing" = can/could/may/might+V, 'You've a chance to be getting Mary and all them today'

BH Inclusive total of verb groups.
I include in this category structures comprising only {V} from {M. Aux. V} structures. Thus 'I <u>wanted</u> to <u>stay</u> and <u>talk</u>' is taken to contain three verb groups. Gerundials also count here. The subset of BH which is finite is given by CT below.

BJ Inclusive total of clauses.
A clause is taken to be a S P {O/C}{OC}({A}) structure, where non-parenthesised elements are obligatory and ordering is not material. I exclude material which I classify as 'fragments' (cf. DH below), and clause initial {S} deletion (cf. CW below), when the latter is not progressively or retrogressively presupposed.
E.g. '<u>What I</u> - Did you see the film?'
DH
'_____ Missed the bus again'
CW
'____went out to work and she hasn't been back since'
CW & BJ

BK No. of localised negative verb forms (excluding tags).

Includes 'double negatives' (he can't not do it 'he can't do it') and never used both as the sole negative operator (I never saw it 'I didn't see it') and as an intensifier (I didn't never do it 'I didn't do it at all'). Cf. also 'You never usually get as many black ones [i.e. Rowntrees Fruit Gums] as that'.

KA Total structures: Pron. + {Aux/M} + Neg + V

KB Total structures as above , where {Aux/M} contracts onto pron., Neg. uncontracted, as in:

E.g. (L) He's not come	(L) He'd not come
(NL) He hasn't come	(NL) He hadn't come/ He wouldn't come

BL Total negative verb groups (excluding clauses with hardly, scarcely, and, for NL varieties, never).

BM No. of localised negative tags (cf. Note (c), p. 210 . Appendix B, above .) - i.e. strictly grammatical tags, excluding e.g. you know, I think, etc.)

375 $(BF^1 + BF^2/BJ)100$

376 (BG/BH)100

377 (BK/BL)100

378 (BM/BN)100

442 (KB/KA)100

BP No. localised question pattern DO --> BE

E.g. (L) What is it he wants?
(NL) What does he want?
(L) Bacon's what it is he wants.
(NL) Bacon's what he wants.

BQ No. of wh-question clauses (and wh- embedded clauses).

379 (BP/BQ)100

QA No. of structures of the type {+Question (inversion/wh-); + Aux; +Neg.}

QB No. of such structures where Neg. = [nɒt] 'not' i.e. uncontracted.

E.g. (L) Does he not do it?
(NL) Doesn't he do it?
(L) What will he not do?
(NL) What won't he do?

443 (QB/QA)100

BR No. of localised phrasal verbs realised.

380 (BR/BH)100 [BH is no. of verb groups.]
E.g. Light the fire on; When did it happen you; I've got money belonging him; He's used with it now; Wait of/on us (=wait for).

BS Total no. of nominal groups.
I define nominal group in terms of actual or potential internal structure rather than surface level function on the practical ground that most of the criteria using nominal group as base are designed to elucidate the distribution of variable internal structure. Essentially this reduces to the usual set of constraints as {± premodifier} HEAD {± postmodifier}.

BT No. nominal groups with localised word order
E.g. 'All the caravan sites are good and I've stayed on them nearly all' [='on nearly all of them']; 'They should sell cup of teas here'; ('I'm going to the clinic for him a box of food' ['for'='to get']); 'I've got a spare flat tyre' [='flat spare tyre' - not a joke] ('There's more parts of Gateshead I would like to live' [='There's parts of Gateshead I'd like to live in more [sc. than this part]'])

381 (BT/BS)100

BU No. of localised pronoun forms.

Certain morphological asymmetries between the L and NL pronoun systems, about which quantitative information is lacking, need to be investigated. The situation is further complicated in L varieties because there are asymmetries between the phonologically weak and strong systems (cf. also criteria 428 & ff. in Appendix A above, p.72 & ff.).

The table which follows presents the differences that I know about from a series of informal and unsystematic observations of speakers who were overheard. It would be an error to imagine that the system given in the table for L varieties was one which all L-variety speakers adhered to in an invariable manner.

		Non-localised						Localised[T]					
Pers.		1st		2nd		3rd		1st		2nd		3rd	
		str.	wk.	str.	wk.	str.	wk.	str.	wk.	str.	wk.	str.	wk.
Sg.	Nom.	[aɪ]	[a(ɪ)]	[jɷu̯]	[jɷ]	[hɪi]	[hɪ]	[ã:]	[ã]	[ji:]	[jə]	[hɪi]	[i]
						[ʃɪi]	[ʃɪ]					[ʃɪi]	[ʃə]
	Acc.	[mɪi]	[mɪ]	[jɷu̯]	[jɷ]	[hɪm]	[ɪm]	[mɪi]	[əz]	[ji:]	[jə]	[hɪm]	[m̩]
										[jɛ^{ə}]	[jəz]		
										[jɪi]			
						[hɜ:]	[hə]					[hø]	[ɛ]
												[hə:]	
												[ha]	
	Gen.	[maɪ]	[mə]	[jɔ(ə)]	[jə]	[hɪz]	[ɪ̣z]	[maɪ]	[mi]	[jɔ:]	[jɛ]	[hɪz]	[ɪz]
										[ja]		[hɪiz]	
						[hɜ:z]	[ɜz]					[høz]	[ɛz]
												[hə:z]	
Pl.	Nom.	[wɪi]	[wɪ]	[jɷu̯]	[jɷ]	[ðeɪ]	[ðɪ]	[wɪi]	[wə]	[ji:^{z}]	[jə^{z}]	[ðe^{ə}]	[ðə]
								[ɤʐ][1]				[jɪiz]	
	Acc.	[ʌs]	[əs]	[jɷu̯]	[jɷ]	[ðɛm]	[ðm̩]	[ɤʐ]	[wə]	[ji:^{z}]	[jə^{z}]	[ðɛm]	[ðm̩]
										[jɪiz]			
	Gen.	[aɷə]	[a]	[jɔ^{ə}]	[jə]	[ðɛə]	[ðə]	[ɛɷə]	[ɛ]	[jɔ:]	[jɛ]	[ðɛ:]	[ðɛ]
								[wɔ]	[wə]	[ja]		[ðɔ]	
									[wɛ]				

1. Note that the degree of voicing in [ɤʐ] varies amongst different speakers from fully voiced to partially voiced.

It is clear from this account that $_L$T varieties have a more independent system of weak forms. In particular the strong/weak [ɤʐ̥]/[wə] and [mɯi]/[əz] pairs cannot be related through any purely phonological derivation. There is, nevertheless, an important alternation between [ɔ] and [ɛ] in the strong and weak forms respectively – [hɔ]/[ɛ], [jɔ]/[jɛ], [wɔ]/[wɛ], [ð̆ɔ]/[ð̆ɛ] – even though many $_L$T variety speakers have lost the strong form from which the weak one could be thought of as deriving. There does not seem to be any reason why a derived form should not in time become a lexical representation in its own right and this can be the only explanation in the case of a speaker who has no form from which the 'derived' form comes.

The most dramatic asymmetry between the NL and $_L$T systems is brought about by two neutralisations in the latter in respect of the NL contrasts. The first is complex in that it consists of neutralisation on two separate levels, one of strong forms , one of weak:
(i) the strong forms of the pl., 1st. pers., nom. and of the pl., 1st. pers., acc. are the same, namely [ɤʐ̥];
(ii) there is a very considerable similarity between this form and the weak form of the sg., 1st. pers., acc., namely [əz].
The second neutralisation is that between the weak forms of the pl., 1st. pers., nom., and the pl., 1st. pers., acc., and the pl., 1st. pers., gen.

382	(BU/BS)100 [BS is the no. of nominal groups.]
BV	No. relative clauses.
BW	No. localised relative clause operators.
BX	No. relative clauses with other differences.

Some $_L$T varieties have what looks like a relative operator as a clause <u>linker</u>, i.e. with a very weak, if indeed any, function inside the 'subordinate' clause. E.g. 'He was going to give her a gas fire <u>which</u> she has a one already'; [and in answer to the question 'What are you looking for?'] 'A list which I don't know what's happened to'; and apparently similar 'Which I've always said, they should have a bus stop there', although this last example may be a representative of another category, i.e. relative clauses without antecedents, e.g. 'Who's got it now, they don't feed it hardly' ['who' =? 'the person . .']; in the previous example 'which' =? 'something'.

On the other hand there are patterns of usage of Ø relative operator which are distinctive: 'That's the only place Ø I've come in contact with people that I haven't really cared for', 'There's still a few engineering firms Ø does that', 'I have one Ø lives in Springwell', 'It was one of those places Ø you weren't on the same type of job', 'There's not much Ø wants done', 'I've got a grandson Ø goes to St. Cuthbert's Grammar School', 'There is places Ø they're different' – but recovered here by use of subject pronoun. Non-fronting of object relative operator: 'I can't talk to somebody I don't know who it is' [= 'I can't talk to somebody who I don't know']. Odd operators: 'A place where you're not used to'.

An interesting set of examples whose patterns partially overlap with those represented here was collected by Anthea Shields (1975) from NL speakers on Radio 4. Her data, as well as my own informal observations, indicate that many of these "non-standardisms" are much more evenly distributed across the whole social spectrum than is usually thought. E.g. "(An organisation should be set up to deal with complaints about X, and Y, and Z and) about motorway caffs **which** I shouldn't think there isn't a person here who doesn't think they're a disgrace to the country" (RP and otherwise 'standard' speaker, Any Questions? Radio 4).

FC No. of restrictive relative clauses with personal/non-personal antecedents where the relative is not the subject of the relative clause 'There was all these bottles of beer what we had bought'

FD Ditto, where relative pronoun = that

FE Ditto, where relative pronoun = Ø

These counters will therefore be able to account for such variations as 'The man who/that/Ø I saw. . .', or 'The car which/that/Ø they stole. . .'

444 (FD/FC)100

445 (FE/FC)100

383 (BW/BV)100

384 (BX/BV)100

BY No. of S-type concords with corporates

BZ No. of P-type concords with corporates
'Corporates' are [+human] collectives. The first concord mark is the only one of concern in this criterion. Thus singular first mark (S-type) 'The government is considering. . .' and plural first mark (P-type) 'The crowd are on its feet. . .' (Nixon 1970).

385 (BZ/BZ+BY)100

FJ No. of localised time adjuncts
E.g. of a Sunday 'on a Sunday' [i.e. 'every Sunday'], not while ten 'not until ten (o'clock)', I'll see him a'=[ə] Saturday '. .this Saturday'.
If [ə] here is a reduced form of on, which is probable, [it could conceivably be a form of of but the point here remains the same] - then it might be argued that there should be a PDV//[ə]// in OU {ɒ}. My justification for **not** having such a PDV is that, as far as can be ascertained its occurrence is restricted to just this time-adjunct. We do not find *[ə]the table for 'on the table', etc.

CB Total of conditioning clauses
Conditioning clauses are separated from linked clauses on the basis of a strong semantic connection between conditioned and conditioning clauses usually involving some modality in the conditioned clause e.g. 'We can do it when he gets here'.

CA No. conditioning clauses introduced by localised clause operators. E.g. 'He can't come being as he's working'; 'With the wife being ill, I'll have to look after her'.

387 (CA/CB)100

In the next set of criteria (388 - 426), the idea is to examine certain features of grammatical complexity and fluency. Though it is clear that the notion of surface grammatical complexity is extremely tricky, either to define or to measure, it has often been assumed that there is some dependent connection between these 'variables' and some socio-psychological factor, or factors. In what follows I make no a priori assumption about either the fact or the likelihood of a connection between the distribution of these criteria in the population and the distribution of social or psychological variables. If there is a dependency then it will express itself in the groupings which are produced by the classificatory process.

388 ((No. of words)/ BJ) - mean length, in words, of clauses.
[BJ is total no. of clauses.]
Clearly, the upper bound of this criterion is undefined.
389 ((BJ-CB)/BJ)100 - % of clauses which are non-conditioning
390 (CB/BJ)100 - % of clauses which are conditioning
CD No. of clause linkers (cf. CB, CA). Exclude from this count SP+SP ('he lived and worked there'), because of CV below.
391 (CD/BJ)100
CE No. of non-conditioning clauses beginning with Nominal group at S. (This and the next three criteria examine the ordering of clause elements in non-conditioning clauses.)
392 (CE/BJ-CB)100
CF No. of non-conditioning clauses beginning with a verbal group (excluding CW types, see below, and necessarily inverted structures). E.g. 'Used to be nice, this place.' 'Turning round, he was.'
393 (CF/BJ - CB)100
CG No. of non-conditioning clauses beginning with an adjunct (excluding 'well', 'so', 'then' and disjunct-type emphatics such as really in '*RF Really * the |way she goes RF on' etc., see DB below). E.g. 'Up the street he goes every morning'.
394 (CG/BJ - CB)100
CH No. of non-conditioning clauses beginning with Object/ Complement (usually marked for theme).
E.g. 'The house I've bought but the rest's on HP'.
395 (CH/BJ - CB)100
CJ No. of nominal group heads = pronoun.
396 (CJ/BS)100
[BS is total no. of nominal groups.]
CK No. nominal groups with 1 premodifying epithet or classifier.

CL No. ditto with 2 ditto.
CM No. ditto with 3 ditto.
Only epithets and classifiers are included on the grounds that the rest of the premodifying items in N.Gps. are less open to affective choice.
397 (CK/BS)100
398 (CL/BS)100
399 (CM/BS)100
CN No. nominal groups with 1 qualifier/postmodifier.
CP No. ditto with 2 ditto.
Here we make no restriction on the kinds of qualifiers since affective choice seems much freer.
400 (CN/BS)100
401 (CP/BS)100
CQ No. N.Gps. with rank-shifted modifiers.
CR No. N.Gps. with rank-shifted qualifiers.
E.g. 'That little-girl (CQ) look of her's (CR)' (Halliday 1961).
402 (CQ/BS)100
403 (CR/BS)100
CS No. co-ordinated N. Gps. [N+N counts as two].
404 (CS/BS)100
CT No. finite V. Gps. in non-fragmentary (see below), non-minor clauses.
CU No. passive V. Gps.
[Excluding 'obligatory passives' e.g. 'I was born in . . .']
CV No. co-ordinated V.Gps. [P+P counts as 2, cf. CS, CD, above.]
405 (CU/CT)100
406 (BH - CT/BH)100 [BH is total V. Gps.]
CW No. of non-turbulent clause initial deletions.
CX No. of non-turbulent 'a', 'the' deletions.
['Non-turbulent' (cf. DH ff. below) refers to the absence of fragments which are intended by the speaker to be deleted by the hearer, hesitation phenomena, etc.]
E.g. 'Wife went to work and she hasn't been back since'; 'Man in the crowd fainted'; 'Car broke down yesterday'; 'Sun never shines'.
408 (CW/BJ - CB)100 Deletions % non-conditioning clauses.

409 (CX/BS)100
CY No. voiced pause [ɛ (±nasal) (±length)]
CZ No. voiced pause [ɜ (±nasal) (±length)]
DA No. voiced pause [ɑ /ʌ (±nasal) (±length)]
Here I am characterising the significance of the phonetic vowel quality of hesitation noises. Typical Tyneside localised is CY, typical NL is CZ.
410 (CY/CY+CZ+DA)100
411 (CZ/CY+CZ+DA)100
412 (DA/CY+CZ+DA)100
DB No. voiced pause plus filler, clause initially.
E.g. well, why [weɩ], I mean, by, you know. These are distinguished from their homophones in non-filler functions.
DC No. voiced pause plus filler, clause finally.
E.g. like, and that, why [weɩ].
Some of these filler items are, in clause-final position, classified as tags (BN), in particular you know, you see. They share with 'central' tags an underlying question form, and the fact that they very often carry nuclear tone, which, in turn, is very often a subordinate tone unit (sensu Crystal 1969).
413 (DB/BJ)100 [BJ is total number of clauses.]
414 (DC/BJ)100
DE No. emphatic-phatic 'I said' plus variants.
E.g. ' 'So I said to him I says' she says ' . . .
DF No. fillers not included under DB, DC, DE.
415 (DE/BJ)100
416 (DF/BJ)100
418 (CY+CZ+DA/BJ)100
DH No. of fragments
[Fragments may be defined as elements at any level of structure whose deletion by the hearer from the message to be decoded is signalled by the speaker in some prosodic and/or paralinguistic fashion.]
419 (DH/BJ)100
DJ No. of repetitions
[This includes all repetitions at any level of clause structure

except rhetorical anacoluthon, which – not surprisingly perhaps – I have never heard in informal conversation.]

420 (DJ/BJ)100

DK No. of repetitions clause initially

DL No. of repetitions of more than 1 institutional word

421 (DK/BJ)100

422 (DL/BJ)100

DM No. of substitutions

DN No. of non-immediate substitutions

[I want to distinguish between immediate substitutions, which are also usually explicit, e.g. 'the littler boy (I mean) **girl**', and non-immediate ones which incorporate some repeated material and which are usually **not** explicit but are very often marked prosodically e.g.

*the little F boy [?] *the little F$^{\text{wider}}$ girl fell over

The tone in the intended string usually copies that in the rejected string and is wider than it; the strings are often separated by a glottal stop; the intended string often has a higher onset and is often louder than the rejected string.

Only the latter, non-immediate substitutions, are counted in DN.]

423 (DM/BJ)100

424 (DN/DM)100

DP No. of expanded substitutions

E.g. 'The point the most important point'...

425 (DP/DM)100

DQ No. of substitutions greater than 3 words

426 (DQ/DM)100

DR No. of weak realisations of 'the' (from a random choice of 25 occurrences of the context) realised as [ði̞] /____** V...

427 (DR) x 4

[Historical note: In the original list of criteria (Pellowe et al. 1972a: 51 ff.), criteria 428-431 were defined in this, their number order place. In this present account however, they are dealt with in Appendix A, above

(p. 72 & ff.) together with other criteria - 194 and following - which account for related processes of reduction/non-reduction in weak (unstressed) forms.]

Criteria 432-438 are concerned with the item yes. This item is of common occurrence (in interviews) and its realisation appears to be important for variety. One sets up a non-exhaustive matrix of variants and their realisations and then monitors the extent to which this fails to account for all the variation met with.

432 "PDV" //jɛs//
States (5): [jɛs] [jɛə] [jʌp] [jʌ] [jɛp]

434 "PDV" //NASAL//
States (2): [(mʔ)(nʔ)] [(mː)(nː)]

436 "PDV" //ai//
States (4): [ɑɩ] [eɩ] [ɑ̃ː] [ɑː]

438 "PDV" //oho//
States (6):[ɒɦ̥ɒ] [ɛɦ̥ɛ] [øɦ̥ø] [əɦ̥ə] [ɑɦ̥ɑ] [mɦ̥m]

4380 "PDV" //OK//
States (3): [əɷkɛɩ] [əɷk±ʔ] [ʔkɛɩ]

Note: The variations in the actual phonetic realisations of these items are handled in the {OU}//PDV//[State] systems above (Appx. A): thus, both [əɷkeɩ] and [ɔkiɩ] are coded here as 4380 state (1).

Appendix D

Strategies and problems in taximetry

<u>Note.</u>
Following Johnson (1970:222), I use "taximetry" rather than the more usual formatives "taxometrics" or "taxonometrics", because, as he says:

> "Taximetrics" is etymologically the best-formed since the first root is the Greek ταξισ, ταξεοσ (or Ionic ταξιοσ) - "an arrangement", of which the combining form in Greek compounds was ταξι- (Liddell & Scott 1864). The word "taxonomy" is badly formed (cf. the French "taxinomie") but is now firmly established by usage. One may hope that "taximetrics" as used by Rogers (1963), will prevail, or perhaps it is not yet too late to substitute the simpler "taximetry", with the ending on the model of "geometry", "trigonometry", "biometry", "anthropometry", etc. Further development of the first of these could lead to the delightful consequence of numerical taxonomists being called "taximeters".

0. Introductory.

Taximetry, classification, and grouping, and the methods which appertain to them, are closely related to the very ancient idea of **types**, whose intellectual roots may be found in Platonic essentialism. (The reality behind "observable", apparently real, nature is in the essences of objects.) Becker (1950) and Lazarsfeld (1962) trace the use of types in the social sciences from ancient times through Weber to the present.

Properties form a property space. Any given typology represents a selected sector of this property space: the selection is made possible by some process of reduction. Lazarsfeld (1962: 467-470) points out that much of one's work requires one to translate concepts into operational instruments which permit the classification of individuals or groups. The investigator only starts out with a vague imagery, and wants to end up with an ordering which does not exist in advance, but is an intended classification. Whatever the outcome, we know that there is only ever a probabilistic relationship between the intended classification and the occurrence of the properties upon which it is based. And since this relationship is at the centre of our interest, the notion of a class is actually less central than that of an ordering. Thus Black (1956) constructs a system of defining type-membership as deviations "from a clear case". He notes, at the same time, that the criteria which would permit this would be very numerous, that they would admit of variation in the degree to which they were met, and that they would behave such that no simple conjunctive or disjunctive combination of them was necessary and sufficient. This is "demanded by the complexity and variability of the phenomena to be described".

Sokal, in an important paper (1962), warns against the advisability of the total rejection of previous theories, or their castigation as absurd, on the grounds that "new thinking" is often (I would prefer to say 'always'), dependent on the old. [N]

[N] If Beckett is to be believed, the writer of Ecclesiastes (1: 9) was in error as to the domain of non-novelty, "The sun, having no alternative, shone on the nothing new" (1938:1, emphasis added).

Goethe, in his anatomical work, laid the foundations of modern typology: anatomical variations should be compared with, or seen as deviating from, the ideal type. He doesn't say how this is to be found, and to expect that he would is to be thinking about types in the wrong manner. The concept of a type will itself show us how such a type is to be found. "Experience must teach us the parts which are common to all animals and also wherein these parts are different in different animals; thence we proceed to abstraction, involving an ordering of these parts and the erection of a general image" (Sokal's (1962) translation of Goethe).

Kalin (1945 - quoted by Sokal (1962)) defines a type as

> the ideal construction of a form from which all separate forms within the category being considered can be thought to be derived,

and again as

> the basic form of that group obtained by abstraction from the subordinated categories and in the final analysis from all relevant individuals. These latter are related to a type as individual cases to a law or as musical variations to the theme of a melody (emphases added).

Most present-day taxonomists eschew typology because of its idealist/essentialist character. The metaphysical basis of types being repugnant to their pretension of an empirical, cause-and-effect manner of constructing classes or groups. Such a puritanical stance is both nugatory and self-deceptive. Witness, for instance, Needham (1965b) who, in tracing some of the sources of methodological difficulty in taxonomy, points out that a good deal of the problem arises from people having an extremely geometrical view of what a classification should look like. He goes on to say that a classification of data should be seen as an epitome of that data, such that originally similar objects behave similarly in the epitome; but he emphasises that the defining characteristics of objects in the epitome may be wholly different from those of the original objects. Surely 'epitome' here is a crypto-typological term ?

(I return to these matters below in section 5.)

* * * * * *

The application of numerical methods and computers to classificatory or taxonomic problems is a recent and rapidly developing field. It is complicated by two factors:

(1) the variability of methods which is a consequence of the different purposes or beliefs of classifiers;

(2) the derivative, and hence not reliable, nature of the statistics of significance and similarity applied in the unusual case of arbitrarily

delimited groups.

This arbitrariness is a reflection of the classifier's purpose, which is itself something outside ' the phenomena themselves', and therefore distorts, in unknowable ways, the 'natural' population from which those phenomena come.

In other words, both the choosing of a method and the evaluation of that method do not rest on well-defined principles. Needham (1965b) emphasises that where automatic classification is really needed there is commonly little or no information on which to base a strong classificatory model.

In this Appendix I draw on the major work in the taximetric field, setting out the major choices available to the investigator at various stages of the classificatory process, pointing out the uncertainties in and consequences of each, and indicating reasons why some set of choices rather than another <u>may</u> be more appropriate for data from language variation.

In all classificatory processes there are essentially four groups of activities associated with: (1) data and criteria, (2) similarity or resemblance (and their converses), (3) grouping or clustering, and (4) diagnosis or allocation. I will deal with each of these in turn.

1. Data and criteria.

Following Sokal & Sneath (1963; and Sneath & Sokal 1973) and others, I shall refer to the objects which should finish up as group members as operational taxonomic units (OTUs). As a useful shorthand we may say that, for the purpose of the Tyneside Linguistic Survey, the OTUs are speakers.[N]

[N] But <u>strictly</u>, the OTUs must be $_aV$s, since any given speaker will usually have more than one $_aV$, though more than one $_aV$ does not usually appear within a single homogeneous context of situation (if such a thing exists). For more discussion see above, Ch.3.

There are few generalisations which can usefully be made about OTUs except that, whatever the heterogeneity of the classification, the sampling of OTUs should be such as to give reasonable probabilities that most groups which are formed will have proportionally representative numbers of OTUs at any given level of similarity (except perhaps the highest few percentage points). Ensuring that this will be the case for linguistic data when some of the varieties are exteremely rare will always be difficult, but I have dealt with this above (cf. Ch. 3 above).

In general this requirement will be difficult to satisfy in any study, whether linguistic or otherwise, since the sample of OTUs is drawn before the groups - for which we wish to establish proportional representation - are formed. Clusters are extremely sensitive to the sampling method used to obtain OTUs, but when we are sampling for those OTUs we do not have the benefit of the clusters to guide our sampling method. (This is the kind of strapless catch we shall become used to.) Sibson (1972: 313) warns of a limitation on numbers of OTUs which arises from computation time:

"60 objects [sc. **N**] is a convenient figure to keep in mind . . .[since] many enumerative methods which at first sight look quite attractive turn out to have a time-dependence of **N!** [i.e. 60 x 59 x 58 x 57 x , . . . , x 1.]."

There are both more, and more important, generalisations to be made about criteria. We must consider:

(a) How many criteria ? What number of criteria will constitute a stable sample from the whole universe of criteria ? ;

(b) What kind of criteria ? Should the sample be drawn from all possible sub-universes or can one manage with fewer than that ?;

(c) What relationships hold between any pair of criteria ? Does one always imply another ?;

(d) What weighting should criteria be given ? When we think one criterion 'more important' than another, what should we do about it ?;

(e) How should we score the different values of a criterion ? Does our similarity measure impose restrictions on the number of values or states which a criterion can have ? Do we want to count mutual absence of a feature in two speakers as a similarity between them ?

The nature of, and the rationale for, the answers to these questions are the most crucial of all in the taximetric process, since not only do they determine the relative levels of similarity for the study as a whole, but also the varying degrees of statistical significance of each coefficient of resemblance between each pair of OTUs. Hence these choices constrain and direct the whole of the inferential process through clustering, texture (Cattell et al. 1966), trees (Sokal & Sneath 1963), and the partitioning of diagnostic keys (Beers & Lockhart 1962).

<u>The number of criteria.</u>

It is accepted as axiomatic by a great majority of workers in computational taximetry that groupings should be based on multiple criteria (Beers et al. 1962; Beers & Lockhart 1962; Carvell & Svartvik 1969; Cattell 1966; Needham 1961b; Sneath & Sokal 1973; Rogers & Tanimoto 1960). And yet, as far as I know, noone has addressed the conflict between this requirement and the general principle of parsimony attributed to Occam: entia non sunt multiplicanda praeter necessitatem. But that things (entia) should not be multiplied beyond necessity surely applies just as importantly to **data** as it does to analytical categories or concepts or hypotheses (D,C1). [N]

[N] Gilbert & Mulkay (1984: 93) render the Razor as "the right theory [is] the one which rationalises the evidence most economically", not seeming to realise that nature does not conveniently leave the evidence in separate heaps-to-be-rationalised in advance of anyone's having a theory. For a theory in which Occam has no place - the razor cuts itself, or turns to water - see Gribbin (1985) & above, Chapter 4.

Needham (1961a) and Good (1962), particularly, have stressed that the multiplicity of criteria must be constrained by relevance. Relevance, that is, to the desired nature of the groups. Thus, to take a simple example, a criterion 'informant has false teeth' - coded positive or negative - has no relevance for the groups desired in this study, though it clearly might be relevant to some other orientation to speech differences. Gower (1969) says that some criteria may not contain any useful classificatory information (e.g. fur length in distinguishing cats from dogs). To be sure, the values of irrelevant criteria vary from OTU to OTU but they are not correlated with the values of other criteria. The inclusion of irrelevant criteria is an implicit form of subjectivity. Johnson (1970) claims that of four important decisions - all of which are subjective - which must be made in choosing a set of methods for classifying, one is the selection of the domain of attributes which we consider relevant to our interest in the objects.

Not only is there no direct and simple way of determining the relevance of criteria for a given study, there is also no way of determining the requisite number of criteria for that study. For instance the number of criteria cannot be determined either from the number of OTUs to be sampled, or from any estimate of the heterogeneity of that sample. On the assumption that 'similarity' is parametric and that it can be sampled, it is possible to make an inference from the graph of confidence limits of different values of similarity coefficients (correlations, associations, distances). For different levels of similarity (ranging from 0 to 1) the curves have different properties. That is, the same confidence limits can be expressed with fewer criteria at a high similarity than at a low one. Figure D1 reproduces a figure from Sokal and Sneath (1963) which plots the 95% confidence limits for similarities of 0.9 and 0.6 against the number of characters on which they were based.

The general inference to be drawn is that unless the number of OTUs is very small (say < 30), and certainly when the number of OTUs is large (say > 300), the minimum number of criteria which ensures stability for the comparison function (e.g. similarity coefficient) is not less than around 100.

It is of the highest importance, however, to note Johnson's (1970) claim that **similarity can never be parametric.** That is, that the assumption of Sokal & Sneath (1963) in plotting a graph like Fig. D1 is untenable, and that consequently inferences derived from it are certain to be erroneous. Gower (1969: 363), surely with tongue in cheek, remarks that the rationale for a maximum of characters is that "there is some "true" value of similarity between pairs of individuals".

Others have suggested not fewer than 60 criteria. Cattell, using 'macrocriteria' which were structurally as well as analytically independent (see the section below on 'Criterial relationships'), and using a coefficient showing similarity of relations as well as of fundaments (see below: 'Similarity') obtained a useful typology with only 12 criteria (Cattell & Coulter 1966; Cattell et al. 1966). Carvell & Svartvik (1969) used more than 70 criteria to classify 146 prepositional strings.

The impossibility of being prescriptive about either the numbers or the domains of the criteria which should be chosen is well reflected in these remarks of Johnson's (1970: 213-4):

> Every "object" ... has an infinitude [emphasis added] of attributes. That is to say, the object itself [emphasis added] does not set any bound beyond which we can say "no further attibutes exist"; this conclusion is not affected [emphasis added] by the practical limitations of our thought ... Any object bears various relationships, tenuous as they may be, to every other physical object or collection of objects in the universe at every point on the world-line of every particle in space-time. There are likewise relationships to the past and future states of the object itself, and indeed to innumerable abstract concepts... There is no a priori reason to stop at any point in

> our search for further attributes. It may be objected that relations to other objects are not intrinsic attributes of the object under consideration . . . [But] all describable characteristics are relations to other objects or concepts; we simply cannot speak meaningfully [emphasis added] of the properties of a thing in itself. . . [added to which] attributes themselves are infinitely [emphasis added] divisible.

Kinds of criteria.

Unfortunately, there has been little systematic work on the consequences of a selection of all or most of the criteria from one system or structure of the OTUs, rather than from another system or structure of the OTUs. That is, even in the biological sciences, it is not known with any generality how the sampling of criteria affects either measures of similarity or the products of different clustering techniques. (Sokal and Rohlf (1966) made a useful start on this problem.)

In linguistic classifications the criteria have most frequently been exclusively morphological or exclusively phonological. The urgent need for sampling experiments of the type implied here seems not to have been realised in any linguistic typological or classificatory studies.

Minkoff (1965) has examined the effects of changing a small number of criteria. The changes he tried were either (a) the definition of their states, (b) the ordering of their states, or (c) withdrawing some states.

Such changes appear to have far reaching effects upon the stability of some similarity coefficients. Johnson (1970) says that one of the important subjective decisions we make in the classificatory process is the "fineness" with which we analyse the features into states. (See above Appendix A and the rationale given there for the hierarchy of {OU}://PDV//: [state].)

Sokal & Sneath (1963) have advanced a pair of related hypotheses in this connection. They are amenable to numerical test and in the meantime are suggestive.

First, they hypothesise that every criterion to be used in the classification is related to more than one underlying factor, and conversely that every underlying factor affects more than one surface-level criterion. There is thus a complex nexus of cause and effect between atomistic criteria and the organs or systems or structures of the OTUs which they are to represent.

Secondly, it is hypothesised that no large classes of factors affect exclusively one class of criteria.

The implications of these hypotheses are that one should utilise criteria from as many as possible of the systems and structures of the OTUs, and that one should use as criteria all features which vary within the sample of OTUs. There are good statistical reasons for rejecting invariant criteria. The most important of these is:

either (a), assuming that all OTUs possessed the invariant criterion, then the similarity coefficients between all pairs of OTUs would be spuriously increased by such criteria, with the result that between-group discrimination would be erroneously decreased and within-group cohesion would be erroneously increased;

or (b), assuming that all OTUs did not possess the invariant criterion, then the similarity coefficients between all pairs of OTUs would be spuriously decreased by such criteria, with the result that between-group discrimination would be erroneously increased and within-group cohesion would be erroneously decreased.

The notion of 'invariant criterion' is not a straightforward one however. In studies of the kind modelled in Chapter 3 above and Appendix B, where a continuously increasing sample is specified, and where various amalgamations and partitionings of samples are necessary to establish measures of stability, a criterion which is invariant for one sample may not be invariant for another. For this reason, it seems to me safe to include invariant criteria provided that (a) there are not very many (no more perhaps than 1 in 20, but this is of course an arbitrary guess), and (b) that they are removed from account when diagnostic keys are to be constructed (D,C2). Gower's remark (n.d.(a): 8), in the context of maximal predictive classifications, supports my view. He writes "correlated variables should be used to reinforce predictive classifications rather than be eliminated".

In general the "inclusion of all [criteria] the investigator has been able to observe should avoid bias" (Sokal & Sneath 1963:95). Johnson (1970: 216) comes to a similar conclusion but through a different critical framework. He notes that it is an arbitrary (i.e. subjective) restriction that "we may reasonably confine our attention to attributes which show a degree of stability or regularity in the individual objects over the time-range in which we are interested". This fits very well with the preoccupations of these studies, since developing a model for specifying variant communicative codes as wholes demands that as many observable criteria from as many linguistic systems as possible be included.

There remains the problem of establishing the states of a criterion, but Johnson doesn't see the problem in this way. Homologies on elementary attributes (properties, undivided criteria) must be established between the objects, and

> this implies setting up one-to-one (pluri-unique) correspondences, over the object set, between some of the elementary attributes of the objects... The establishment of [this] correspondence ranging over the object set ... is tantamount to a classification of the attributes: we have assigned a certain one of the elementary attributes of each object to a class. From the property by which we define this class we derive a pluri-state attribute applicable to all members of the object set; the elementary attributes then become the individual "states" (1970: 208, emphasis added).

There are two problematic consequences for the property space as a result of such considerations, both of which are unsolved. (a) If some pluri-state attributes depend for their expression on the existence of certain states of others, then the attribute space is non-homogeneous - and this can't be solved by weighting or redefinition of attributes. (b) Some states which may be reasonably grouped as a single attribute do not admit of a serial arrangement, and a property space which incorporated such attributes would have to express some attributes as linear coordinates and others as higher order (triangular, tetrahedral) coordinates, i.e. "dimensions within dimensions".

Criterial relationships.

However, limitations other than those of numbers and structural representativeness must also affect the choice of criteria. First, except in the special case of the acceptability of invariant criteria which I have just discussed, sample invariant criteria must be eliminated. In some cases such criteria may be made variant across the OTUs simply by redefinition.

Let us take a hypothetical example: imagine that we had defined a criterion as 'informant uses [ə]' - which would have had an invariant (positive) response for all informants. If we redefine this criterion as 'infomant uses [ə] in stressed and unstressed syllables', we would obtain variable responses (since NL-varieties do not permit [ə] in stressed syllables). This example shows that invariance of criteria is **not** always a property of the sample of OTUs, but may arise from insufficiently thoughtful definitions.

Secondly, because they would blur distinctions between sets of similarity coefficients, it is necessary to exclude criteria which are logically correlated and hence always imply each other's presence or absence. They might be logically correlated structurally (that is because of some redundancy within or between linguistic systems), or because of the method of analysis [sc. transciption], or because of the method of coding [sc. representation of criterial states]. If the logical correlation between criteria is structural then the problem cannot be solved and one of the criteria must be dropped. But correlation must occur across the whole sample, since exclusion on the grounds of correlation across a subsample would be too strong a condition. There then arises a procedural problem, since we need to have a full analysis of the criteria in the sample before we can decide what criteria to analyse in the sample. (The same Catch in a different place.)

On the other hand, as MacNaughton-Smith shows (1965: 21), if the association between all attributes (considered in pairs) were zero, then most (i.e. polythetic) methods of making clusters would run into serious difficulties, and some (i.e. monothetic) methods would be downright impossible.

Thirdly, it sometimes happens that the speech feature of an OTU to which a criterion would respond is missing for some reason. This is a different state of affairs from when the criterion is negative and we need some indication that the criterion is not applicable. If a criterion is N/A for OTU 1, then comparison on that criterion with OTU 2 is meaningless. It may be scored NC (for 'no comparison'). Clearly, both the case of OTUs having many NC criteria, and the case of criteria scored NC for many OTUs would be a considerable source of difficulty, since they would in general, and in ways not likely to be symmetrical, depress the levels of similarity, and hence make the resultant clusters less easy to interpret.

Weighting criteria.

Apart from the requirement of multiple criteria, the axiom that criteria must be given equal nominal weight ab initio is the strongest and most widely agreed requirement amongst computational taximeters. This insistence derives ultimately from the 18th century botanist M.

Adanson [(1763) **Familles des plantes. Vol I**. Paris: Vincent]. Cogent accounts of the elementary logic are widespread (Beers & Lockhart 1962; Carvell & Svartvik 1969; Gilmour 1937; Sokal & Sneath 1963).

Adanson's basic principles were:

(i) classification is to be based on many attributes;

(ii) groupings are to be formed on many equally weighted attributes from as many systems of the objects as possible;

(iii) one is to refuse completely to attach a priori weights to attributes.

But there are dissenters from this view e.g. Proctor & Kendrick (1963), Kendrick (1964), Kendrick & Proctor (1964). As MacNaughton-Smith (1965: 11) writes:

> weighted measures have aroused great objection . . . out of exaggerated respect for. . . Adanson [who] was concerned to propose objective methods, but . . .was not apparently aware of any distinction between weighting as a mathematical device and subjective bias as a scientific failing; there can be no objection in principle to any impersonal weighting procedure if it can be shown to be more efficient than equal weighting.

Goodall (1966) concurs. He points out that the use of the probabilistic similarity coefficient (see below) gives differential weighting according to the commonness or rarity of the values observed. Clearly objections to differential weighting can have little force when the different weights are inherent in the data themselves. Gower (n.d.(b): 11) adds that it is possible to weight an attribute as the result of comparing it in a pair of individuals - rather than intrinsically. Further, he suggests, disagreement between attributes may be held more important/informative than agreement between them.

Cattell (1949; Cattell et al. 1966) emphasises with numerical evidence that this qualification of 'equal weight' is nominal. That is, we may treat the criteria as if they had equal weight, but the factors which underlie them may very well have different loadings. Nevertheless such loadings cannot be investigated until we have generated the groups whose internal and external criterial variance gives evidence of those loadings. (Again, we find the general empiricist catch-22 of one needing to know everything-in-general before one can know something-in-particular (D,C3).)

There is a common assumption that the ab initio weightings which we feel the criteria deserve are collectively equivalent to the diagnostic key which we would otherwise obtain from the groups to be formed. But this is likely, on many occasions to be an erroneous assumption (D,C4), since a diagnostic key does not only, or even primarily, depend upon the number of criteria and the relationships between them, but also depends on the number and diversity of OTUs, the nature of the similarity coefficient, and the method of forming clusters.

As far as I know there is no work in which unweighted criteria have been used to produce a classification of a sample which is then used as a source of weightings of the criteria to be applied to a second comparable sample. The second sample would also have to be analysed in terms of the old unweighted criteria to check the reasonableness of the weightings.[N]

[N] The difficulty about this minimal thought-experiment, of course, is that no second sample is ever comparable enough with another, even if it is one drawn by the same method from the same population as the first ! (A similar problem is discussed in Ch. 3 above, concerning the optimal sampling methods for urban speech surveys; and in Appx. B, Section D, above.)

Cattell & Coulter (1966) do give a variant formula for their coefficient,

r_p, which allows for criterial weights, and certain association coefficients also have variants giving weights (Sokal & Sneath 1963).

Coding criterial responses.

There are three general types of criteria:

(1) dichotomous (qualitative: +/-; quantitative 1/0);

(2) quantitative multistate (0, 0.5, 1.0, 1.5, 2.0,.... or 10, 100, 1000,....etc.);

(3) qualitative multistate (0, 1, 2, 3, 4,......).

The coding of criteria is of extreme importance, since apart from being directly related to the problems of logically correlated, non-comparable and invariant criteria, the method of coding which is adopted can affect the values of individual similarity coefficients and hence the levels at which individuals join clusters.

In the original proposals for bacteriological taximetry (Sneath 1957a, 1957b) coding was by means of '+' (the OTU responds positively to the criterion), '-' (negatively), 'NC' (non-comparably), where each of these was interpreted literally and applied to dichotomies as follows:

		criteria 1	criteria 2
	A	+	+
	B	-	+
OTUs	C	+	-
	D	NC	+
	E	-	-

Here the situation appears to be simple. Under Sneath's (1957b) original suggestion, the similarities amongst these OTUs are AB 1/2, AC 1/2, AD 1/1, AE 0/2, BC 0/2, BD 1/1, BE 0/2, CD 0/1, CE 0/2, DE 0/1. Sneath excluded negative matches as similarities (B-E on criterion 1; C-E on criterion 2) on the grounds that the 'universe of possible negative criteria is infinite'.[N]

[N] There are two objections to this: first, a negative match on a positive criterion need not imply the use of negative criteria; secondly, depending upon the coding conventions and depending upon imagination the universe of positive criteria may not be as finite as we would like to think. (See the remarks of Johnson (1970: 213) quoted above (p.433), in particular, "every object has an infinitude of attributes".)

Further, MacNaughton-Smith (1965: 10) declares that "it is doubtful whether [omitting negative matches] should ever really be done".

The kind of dichotomy (dichotomy$_1$) illustrated above, can be treated satisfactorily in this fashion, but not all dichotomies represent the presence or absence of a feature. There is a type of dichotomy ($dichotomy_2$) which depends upon the presence of one feature or the presence of a different feature. In such cases we must either permit negative matches, or we must use two criteria and some coding convention which avoids the logical negative correlation between them. A $dichotomy_2$ criterion - depending on presence of x or presence of y - can either be represented as:

		criterion (xy)
	A	+
OTUs	B	-
	C	-
	D	+

whereby AD have 100% similarity, and BC have 100% similarity, or as:

		criterion (x)	criterion (y)
	A	+	NC
OTUs	B	NC	+
	C	NC	+
	D	+	NC

whereby A & D have 100% similarity, and B & C have 100% similarity. The differences in the two cases come about from a different convention about the meanings of '+', '-', and 'NC' (but under identical definitions of what constitutes the basis for similarity). In the first case '-' signifies, by implication, that feature **y** is present, in the second case '+' signifies that feature **y** is present. But the second case depends on a non-literal use of 'NC', in which it is being used as a buffer against spurious comparisons.

The same technique may be extended to multistate criteria such that a criterion with **n** states may be reduced to **n**-1 two-state (binary) criteria. If the criterion is quantitative and if it can be assumed that the difference between one state and another represents an additive feature in the loading on an underlying factor, then <u>additive coding</u> can be used. Say we have a 5-state quantitative criterion, **x** (0, 1, 2, 3, 4), then we can express any one of these states by using a suitable number of two-state criteria:

	states	2-state criteria a	b	c	d
	0	-	-	-	-
	1	+	-	-	-
criterion **x**	2	+	+	-	-
	3	+	+	+	-
	4	+	+	+	+

[Thus, **x**(0) can be represented by a(-); **x**(1) by a(+); **x**(2) by a(+) and b(+); x(3) by a(+) and b(+) and c(+), and so on.]

If we do not wish to involve ourselves in the theoretical problems implied for our material by additive coding, then we can use non-additive coding. In this case a criterion z with **n**-states is converted into **n** 2-state criteria. Its logic is as follows.

For two OTUs to have in common <u>some</u> value of a multistate criterion and for them to have the <u>same</u> value of it are two different things. If we call the first of these things X and the second Y and the individuals share X, but not Y, then their similarity is 50%. If one of them has <u>no</u> value for X, then it must be NC for Y, thus:

		2-state criteria				
	states	X	Y1	Y2	Y3	Y4
	0	-	NC	NC	NC	NC
	1	+	+	-	-	-
criterion z	2	+	NC	+	-	-
	3	+	NC	NC	+	-
	4	+	NC	NC	NC	+

These methods of forcing multistate variables to behave as binary variables are fairly well understood. On the other hand, the question of how to dichotomise <u>continuous</u> variables is extremely complicated, as

MacNaughton-Smith says. The setting (1965: 22) of the cut-off for one variable affects the optimum settings of other variables, and a good setting of the cut-offs for the initial group of variables may not be good over the whole analysis.

Negative matches have previously been considered to be theoretically inadmissible in the construction of similarity coefficients, but a more pragmatic attitude towards them seems to me to be useful (D,C5). There may be some criteria for which mutual absence of some state in two OTUs is important, possibly to the extent of its being diagnostic (see below). Furthermore, if negative matches are excluded, the denominator of the expression of similarity

$(S_{i,j} = n_{same}^{i,j} \ / \ n_{same}^{i,j} + n_{different}^{i,j})$

will vary with each pair of OTUs. This in turn implies that the statistical significance of individual S measures will also vary, and this will have more serious consequences in proportion as samples of OTUs are more heterogeneous.

There are thus both practical and statistical reasons for including negative matches in the calculation of similarities (Beers & Lockhart

1962, Beers et al. 1962).

Standardisation of criteria.

When criteria of different kinds - qualitative and quantitative with varying numbers of states - enter into a coefficient of similarity it is necessary to reduce all the states of each one across all the OTUs to a mean of 0 and a standard deviation of 1. This may be done on the raw score matrix, or by pairs of OTUs as the similarity coefficients are calculated.

The mean of any criterion, **c**, is

$$\bar{X}_c = 1/t\,(\sum_{j=1}^{t} (X_{cj}))$$

where t is the total number of OTUs, and j is any OTU.

Its standard deviation is

$$SD_c = [(1/(t-1))\,(\sum_{j=1}^{t} X^2{}_{1,j} - (\sum_{j=1}^{t} X_{1,j})^2/t\ )]^{1/2}\ ;$$

then raw scores are standardised by:

$$X^{c}_{i,j} = X_{i,j} - \bar{X}_{i,j} / SD_{c}\,.$$

Nevertheless, when criteria of different kinds are being used to group OTUs, it has been claimed that the operation of standardisation, even though it determines the distribution of criteria individually, in no way establishes the stability or significance of similarity coefficients. Kendall, for instance, makes this criticism; and then adds that "if we reverse the sign of a given variable, for example (and there is no reason why we should not), we obtain a completely different set of correlations" (1966:181). He proposed to render all analyses distribution-free by using rank correlations. Cattell et al. (1966) express the same sorts of fears.

The criticism seems to be a valid one when the different criteria involve mixed metrics. That is, when different criteria represent things with different scales and/or with different units. However, standardisation of qualitative criteria (whether 2- or multi- state) seems to me to be quite irrelevant for the stability of coefficients. In particular, reversing the sign of such a criterion would have no meaning numerically - though it would in terms of the coding conventions (see above).

For this reason, Cochran and Hopkins (1961) propose that in general if there is a mixture of qualitative and quantitative data, and if the latter are not too numerous, the latter should be turned into the former (see above for the method). But they worry about the amount of discriminatory power which is lost, and what precisely the best partition points will be. They also suggest, in this context, that the probability of misclassification will always be underestimated.

Criteria: summary.

(1) As representative and as large a number of criteria as possible are chosen. N

N On 'representativeness' MacNaughton-Smith writes "[the investigator] may try to achieve objectivity by asking "What information would be included not by myself but by a hypothetical 'ideal researcher'? What objective criteria of relevance and interest can I set up? "... this important and fascinating topic has not yet been tackled" (1965: 21). On the notion 'attribute' [sc. criterion] Watson et al. (1966: 492) declare that it remains undefined, since the OTU is a perfectly integrated unit and there is no reason why it should lend itself to subdivision for our convenience.

---

(2) Criteria must not be logically correlated.

(3) Depending upon the nature and number of their possible states, they are coded in different ways all of which should contribute to the stability of the similarity coefficient, though how this is to be ensured is never actually revealed to us, and seems, in fact, not to be known.

(4) The totality of coded criteria for all the OTUs then forms a feature matrix or a basic data-relations matrix (BDRM - Cattell 1966). N

N Nevertheless, Rijsbergen (1971) shows that there are considerable advantages to be got from storing the data in an ordered tree rather than in a matrix. In particular, further data can be added without requiring total recomputation of the tree.

---

2. Similarity/dissimilarity.

Some of the "properties which adhere to one's own intuitive concept of likeness [are] unfortunately . . . mutually exclusive", nevertheless MacNaughton-Smith (1965: 9) is able to provide us with four desirable mathematical properties of a dissimilarity function. It should:

(a) simultaneously maximise between-set dissimilarity and minimise within-set dissimilarity;

(b) make it certain that some measure will decrease as the hierarchy descends (the information coefficient and Euclidian distance show this property, others don't);

(c) be defined in such a way that its values are comparable at different levels of the hierarchy (i.e. at different stages of the analysis);

(d) not be defined in concession to visualisability.

Furthermore, the notion of a good dissimilarity function is closely related to that of an efficient discriminant [sensu Fisher, see below 'Diagnostic keys'].

The assessment of similarity or relationship is not new to statistical methods. Whichever coefficient is used , and whatever is desired of the final groups, the starting point is always the set of values of **n** variates for **N** OTUs, usually in the form of an (**n** x **N**) matrix (BDRM). Its generalised form is:

Q- OTUs/ ids/ species etc.

		1	2	.....	j	.....	N
	1	1,1	1,2	.....	1,j	.....	1,N
	2	2,1	2,2	.....	2,j	.....	2,N
R-	..						
criteria/	..						
characters/	i	i,1	i,2	.....	i,j	.....	i,N
tests etc.	..						
	..						
	n	n,1	n,2	.....	n,j	.....	n,N

where the intersect i,j contains the relevant criterial state of the i^{th} criterion for the j^{th} OTU .

Using such a matrix one may determine two different sorts of relationships. Either one may group criteria (rows) across all OTUs, or one may group OTUs (columns) down all criteria; the former is called R-technique, the latter Q-technique. Classifiers usually assume (but cf. Cattell et al. 1966) that the methods used to arrive at the relationships between criteria are transferable, without adaptation, to a grouping of OTUs.

There are four main groups of coefficients used for assessing relationship amongst either criteria or OTUs. There are many variants of each group but my discussion of these will be eclectic. The four main groups are (a) correlation coefficients, (b) contingency or association coefficients, (c) distance coefficients (whether Euclidian or non-Euclidian), (d) information coefficients.

Correlation coefficients.

Only the product-moment variant of this coefficient has been in general use in computational taximetry (Sokal & Michener 1958). The data employed usually involve a majority of multi-state criteria. The formula takes moments around the mean, and where quantitative multi-state criteria are concerned this process represents the magnitude of the mismatches between OTUs. Limits of correlation coefficients range from +1 to -1, but what the meaning might be of a complete negative correlation between OTUs in the case of linguistic varieties is difficult to imagine. The major problem with the use of correlation coefficients in work on linguistic variation is that such coefficients usually assume a bivariate normal distribution. Such an assumption is difficult to support when arbitrarily coded data (qualitative multi-state criteria having variable numbers of states) make the vectors of each column in the matrix so heterogeneous.

Cattell et al. (1966) claim that even if the matrix of correlation coefficients is only used for clustering (Q' technique) and is not factored (Q technique), it is still not satisfactory for grouping OTUs. Let us consider a profile of resemblance as consisting of two components - shape and level; then the reason why correlation coefficients do not satisfactorily cluster OTUs is that the formula for correlation coefficients, by cancelling means and Σ s, destroys the level of the profile whether it is the same or different for any pair of OTUs.

Contingency or association coefficients.

Use of contingency coefficients demands that all criteria are two-state ones. (It is immaterial whether they are truly binary (+/-), dichotomous (presence *a*/ presence *b*), or some transformation of multi-state criteria (see above).) The coefficients operate on the properties of a 2x2 table for each pair of OTUs, where upper case subscripts refer to positive and lower case subscripts to negative values for some criterion.

		OTU j +	OTU j −	total
OTU k	+	n_{JK}	n_{jK}	n_K
	−	n_{Jk}	n_{jk}	n_k
	total	n_J	n_j	n

Coefficients vary as to whether or not they give equal or unequal weight to matched pairs (n_{JK}, n_{jk}) or unmatched pairs (n_{Jk}, n_{jK}). More significantly they vary as to whether they include negative matches (n_{jk}) in the numerator , i.e. as contributing to similarities.

The simple forms of these coefficients are thus:

(i) $$A_{jk} = n_{JK} / n_{JK} + n_{jk} + n_{Jk} + n_{jK}$$

(here, negative matches do <u>not</u> count as similarities) and

(ii) $A'_{jk} = n_{JK} + n_{jk} / n_{JK} + n_{jk} + n_{Jk} + n_{jK}$

(here, negative matches do count as similarities), where the distribution of the numerator n_{JK} (or $n_{JK} + n_{jk}$) is normalised by reference to the table of the degrees of freedom associated with the χ^2 median (Cattell et al. 1966; Cattell 1949).

Distance coefficients.

It is possible to conceptualise each criterion upon which the OTUs are to be described as a dimension. The comparison of j OTUs on n criteria will then be represented as a dispersion of j points in an n-dimensional hyperspace. In general most workers in taximetry have assumed that this space is Euclidian , or metric, such that if A is close to B, and B is close to C, then A is close to C (Sokal & Sneath 1963; Cattell 1966; Sneath & Sokal 1973 etc.). Some investigators have suggested that it might on occasion be more relevant to use a semi-metric n-dimensional space such that if A is close to B, and B is close to C, then A may be, but need not necessarily be, close to C (Rogers & Tanimoto 1960; Beers & Lockhart

1962). The same effect is obtained by permitting topological deformations of a contingent cluster boundary (Good 1965a; Needham 1965b). (See below: clustering.)

The two Euclidian treatments of taxonomic distance depend, as one might expect, on an n-dimensional extension of Pythagoras' theorem:

$$d_{i,j} = \sqrt{\sum_{k=1}^{n} (x_{k,i} - x_{k,j})^2}$$

where x is the value of k, any criterion (all assumed to be independent), and i,j is any pair of OTUs.

The first of the two treatments of relationship which assumes Euclidian space simply gives relationship by distance. Thus, a pair of OTUs i,j, having a low value of $d_{i,j}$ are very similar. The serious drawback of this method is that it is known that as the numerical value of the coefficient decreases so does the stability of the fiducial limits (cf. Fig. D1). That is, as similarity increases, our confidence in the trustworthiness of that similarity decreases: surely this is exactly the opposite of our experience as classifying creatures in daily life (D,C6).

The second treatment of relationship by Euclidian distance attempts to correct this by taking the reciprocal of distance. Thus, using the $d_{i,j}$ from the expression above:

$$S_{i,j} = 1/d_{i,j}\,; \qquad S_{i,j} \text{ being the similarity between i and j.}$$

However this treatment also raises a problem which I have touched upon above. This is the problem of a denominator which, even if criteria are standardised, varies radically between one pair of OTUs and the next. Thus there are statistical problems attaching to both

$$S = n_{same} / n_{same} + n_{different}\,, \text{ and}$$

$$D = n_{same} + n_{different} / n_{same}\,.$$

The difficulty can be removed in both cases by use of a logarithmic form of the expression. As I shall point out below this has useful diagnostic properties. It represents the dispersion of the OTUs in a semi-metric space. (Cf. Rogers & Tanimoto 1960; Beers & Lockhart 1962; Lockhart & Hartman 1963.)

Information coefficients.

These coefficients exploit the notion, due, in our times at least, to Shannon, that probability is a very closely related transform of information (cf. Gower 1969). The following account is derived from that of MacNaughton-Smith (1965), but those in Goodall (1966) or Watson et al. (1966) are equally useful.

Let an initial set {A} consist of **n** individuals observed to possess or lack **m** attributes. The row total - the number of attributes possessed by **i** is $x_{.i}$. The column total - the number of individuals possessing **j** is $x_{.j}$.

For every x_{ij} there corresponds a probability, p_{ij}, that $x_{ij} = 1$. It is frequently desirable to define similarity in such a way that in any final set, for any given attribute, j, the quantity p_{ij} is the same for all the individuals in the set. This probability is p_{0j}. The best estimate of p_{0j} is $x_{.j}/n$, the attributes being assumed independent.

Therefore, for any individual, multiply all p_{ij} of its possessed attributes, and all $(1-p_{ij})$ of its lacked attributes; the product is the equivalent of an individual's <u>likelihood</u>, $l_i\{A\}$ - the likelihood of i, if {A} is a final set. If you multiply all $l_i\{A\}$ for individuals in {A} then you get the set likelihood $l_{\{A\}}$, which takes on maximum value in respect of all p_{oj} when the estimate $p_{oj} = x_{.j}/n$ is used.

Then:

$$\log L_{\{A\}} = \sum_{j=1}^{m} (x_{.j}\log x_{.j} + (n-x_{.j})\log(n-x_{.j}) - mn \log n)$$

If one potential division of {A} is into {B} and {C}, consisting of n_1 and n_2 members respectively, and where **x** in {A} is equivalent to **y** in {B} and z in {C}, then $\log L_{\{B\}}$ and $\log L_{\{C\}}$ can be got in the same manner as $\log L_{\{A\}}$.

Then the information statistic can be defined as:

$$INF_{B,C} = \log_e L_{\{B\}} + \log_e L_{\{C\}} - \log_e L_{\{A\}}.$$

(There is also a two-parameter model, in which each individual has an "ability", a_i, and each attribute a "difficulty", $1/b_j$. The probability, p_{ij}, would then be given by: $a_i.b_j/1+a_i.b_j$.)

Given the two way association between two individuals:

		k +	k −
j	+	a_{jk}	b_{jk}
j	−	c_{jk}	d_{jk}

then the information statistic, measuring the association between j and k, is

$$INF_{jk} = a_{jk}\log_e a_{jk} + b_{jk}\log_e b_{jk} + c_{jk}\log_e c_{jk} + d_{jk}\log_e d_{jk} -$$
$$x_{.j}\log_e x_{.j} - (n-x_{.j})\log_e(n-x_{.j}) - x_{.k}\log_e x_{.k} -$$
$$(n-x_{.k})\log_e(n-x_{.k}) + n\log_e n.$$

Goodall (1966) uses a probabilistic similarity index which shows reasonable agreement with the simple matching coefficient of Sneath (1957a). It is defined as the 1-complement of the probability that the two OTUs would have the observed, or a greater, degree of similarity if the attribute values were assorted at random.

Lance & Williams (1966) claim that <u>their</u> information statistic is immune from any failure of monotonicity. A good general initial strategy would therefore be: information statistic and centroid. Let a system be capable of existing in any one of a number of discrete states, and let the probability of the i^{th} state be p_i, then the entropy of the system,

$$H = \Sigma_i (p_i \log p_i) \text{ [Shannon]},$$

and the mean entropy for n individuals having ±s attributes is, given that the probability of the presence of the j^{th} is p_j,

$$H = -\sum_{j=1}^{s} (p_j . \log . p_j + (1-p_j) \log (1-p_j)).$$

If a_j individuals have the j^{th} attribute, then the best estimate of p_j is a_j/n, and the information content of the whole system is $i = nH$. The increase of i on any fusion is to be a minimum;

$$i = sn \log n - \sum_{j=1}^{s} (a_j \log a_j + (n-a_j) \log (n-a_j))$$

(the base of the logarithm is irrelevant).

Finally, Watson et al. (1966) aver that:

> the information statistic has been invaluable [because] it has enabled us to set subdivision levels but . . . it has grouped together peripheral genera irrespective of their affinities . . . a property which would be

> useful . . . in defining the 'central' groups . . . but which could, in relict groups, where outliers are particularly numerous, be a disadvantage . . The value of this statistic . . . may well be to provide a preliminary analysis of degree of heterogeneity (1966: 497).

But note that this assumes that <u>which</u> OTUs are outliers is already **known** !

There are several difficulties involved in defining coefficients of similarity/dissimilarity.

First, and the most intractable, is that there is no parametric value for similarity (Johnson 1970: 216, <u>pace</u> Sokal & Sneath 1963). Moss (1971: 312-3) agrees. The application of conventional tests of statistical significance is complicated in the case of similarity coefficients because they do not ever show conventional distributions - which is of course what significance tests depend on. On the contrary, they always demonstrate kurtosis and skewness which cannot be eliminated by any kind of scaling or conversion - e.g. by using a logarithmic representation. The very fact that the distributions are certainly abnormal is what prompts investigators to turn to taximetric methods.

Secondly, the meaning of any particular similarity coefficient is unclear. Johnson (1970: 214) warns that "there can exist no absolute measure of similarity (i.e. matching correspondence) between non-identical sets of attributes which are infinite, unbounded and unconstrained". Gower (1969) growls - with commendable brevity - that "no meaning, as such, can be given to a level of similarity". And Goodall (1966: 36) emphasises that similarity indices refer to a specified universe of discourse and have no meaning in isolation - "they are indeterminate for a particular pair of [OTUs], until these . . . are considered as two of a larger collection".

Thirdly, the relationship between similarity coefficients and the attributes on which they are formed is not recoverable. Sibson (1972) notes that the expression of a pairwise measure of similarity loses all information about the relationships between the attributes. In ecology, where one is trying to say something simultaneously about objects and attributes, this is clearly unsatisfactory. He also points out that it is safest not to assume that there is any stronger relationship between different values of any comparison function (CF - i.e. correlation, association, distance coefficients) than that of ordination. To this

Williams et al. add:

> it has always been known that these [similarity, hierarchical] methods may fail if very few attributes are available, or if many of the attributes are lacked (or possessed) by nearly all the individuals . . . Thus both Sneath's and Tanimoto's methods may be inappropriate in . . . ecology, where few species may be present, and some of these may be rare; [similarly in] psychology, sociology or criminology. Sneath has always stressed that his methods are not . . . applicable to such data (1964: 426).

Similarity: summary.

Some suitable coefficent of relationship is computed for each pair of OTUs on the basis of the entries in the BDRM.[N]

[N] Sokal & Michener (1967) have investigated the use of different coefficients on the same data proving only that the best choice of coefficient cannot be determined by inspection of the raw data.

These coefficients are entered in a matrix bounded on both sides by an identically ordered list of OTUs. This is the similarity matrix (SIMX).

		OTUs					
		1	2	3	4	..	N
	1		1,2	1,3	1,4	..	1,N
	2	2,1		2,3	2,4	..	
	3	3,1	3,2		3,4	..	
OTUs	4	4,1	4,2	4,3		..	
	..	..	..	..	..		
	..	..	..	..	..	..	
	N						

This matrix is often found in reduced form as a sub-diagonal matrix (triangular) since the similarity (1,2) is the same as the similarity (2,1).

*****

3. Clustering & grouping.

As Sparck-Jones and Jackson laconically have it: **if** similarity is well-understood, then grouping procedures are **ill-understood**. And later they say (1967: 37) that "it is as difficult to infer how a classification as a whole will work as it is to predict the effect of individual classes [sc. upon the remainder]".

The general requirement of a clustering procedure is that it express the contents of the SIMX in an economical, memorable and suggestive fashion. (We should recall, however, that this 'expression' is not without its own price - definitions of similarity involve the assumption that in a final set all attributes are independent of each other (cf. MacNaughton-Smith 1965: 13).)

One way in which clustering can achieve this result is by producing a nested hierarchy, that is by realising a large number of groups at one rank as a smaller number of composite groups at a higher rank. Establishing such a hierarchy depends upon finding a clumped or multimodal distribution of the coefficients in the SIMX. [N]

N The desire to classify things at all implies that the classifier has a strong hunch that the distribution is at least bi-modal (Rohlf & Sokal 1962; Sneath 1957b; Cattell 1966).

But stratified or nested clusterings are not <u>necessarily</u> hierarchical: if they are, then at any given level the clusters are disjoint; if they are not then the clusters may overlap. Non-hierarchic stratified clustering is more likely to preserve the meanings of the comparison functions (CFs) than hierarchic clustering (Jardine & Sibson 1971).

However, Lance & Williams (1966b) claim that hierarchical classification has computational advantages over non-hierarchical. But actually we know from Macnaughton-Smith (1965) that this has not primarily to do with hierarchicisation, but with whether we are successively <u>dividing</u> the population or groups, or whether we are successively <u>fusing</u> the nearest pair of individuals or groups. In the first case computation time is proportional to $[2^{n-1} -1]$; in the second case to only $[1/2n\ (n-1)]$ and rarely to more than $[n(n-1)]$ for a complete analysis. Lance & Williams (1966b) also provide three criteria for assessing the performance of various combinations of strategies:

(a) the value of the CF should change only monotonically (i.e. in the same direction and at a compatible rate) after each fusion;

(b) fusion should apply to separate groups and should not consist in the continual adding of single individuals;

(c) the coefficient should define a level below which further fusion may be regarded as unprofitable. [N]

[N] But it is impossible to see how a coefficient, which is merely a fixed calculation, could define the profitability or utility of a group, which presumably depends on the interaction between the mental constructs of the interpretation of the final groups and of the original purpose (D,C7).

Lance & Williams (1967) appear to contradict their own claims for the computational efficiency of hierarchic arrangements. If clustering optimises within-group homogeneity, then hierarchicising optimises a route from the population to groups (divisive methods) or from groups to the population (fusion methods). The former can operate on a subset of the population (re-allocating where necessary), but clearly the latter must operate on the whole population.

However, Lance & Williams do provide a useful metaphor by which various methods can be categorised. The use of the information statistic (see above) tends to be space dilating. That is, as a cluster grows, it becomes increasingly difficult for that cluster to capture new members; thus, as analysis proceeds new OTUs tend to be allocated to the smallest clusters regardless of attributes. And conversely, the measure of mean intra-cluster inter-element distance is a space contracting measure which gives increasing magnetic force to larger groups. Space conservation is, needless to say, an ideal not likely to be attained !

Multi-dimensional (m-d) scaling (Kruskal 1964 a, b) exploits such problems by fitting the similarity coefficient values to an n-dimensional representation in respect of stress reduction. (Stress is determined by minimising the sum of squares.) This strain, according to Hodson et al. (1966), indicates the degree to which a given configuration fails to reflect actual similarities. In most cases increasing the dimensionality reduces the stress. Sibson (1972) points out that classifying gives a well-defined result on any data, but that m-d scaling cannot be expected to give a unique answer since (a) it tries the hill-climbing process from various starting points and may, for a reasonable stress value get trapped in a local optimum, and (b) it may give widely different answers for the same stress, in which case it may not be the appropriate method. Kendall (1971) wants the available programs for m-d scaling to be modified such that colliding objects can 'pass through' one another (during the iterated perturbations of the successive configurations). He also warns that a large number of equally dissimilar pairs (being scaled on 1 dimension) may show up as a horseshoe when the data are known to be linear.

In respect of all methods it is important to bear in mind Gower's (1969) remark concerning abundance. He is discussing the reasons for the development of the weighted mean-pair group method from the unweighted mean-pair group method. The latter replaces the two nearest neighbours by a single member at their joint centre of gravity (centroid), a location which is clearly sensitive to the relative sizes of the two groups. And though the former avoids this problem, he points out that abundance is not a solved problem in classificatory studies. (See above, Ch. 3, for a characterisation of the opposite problem - the rarity of NL varieties.)

Shading the similarity matrix & dendrograms.

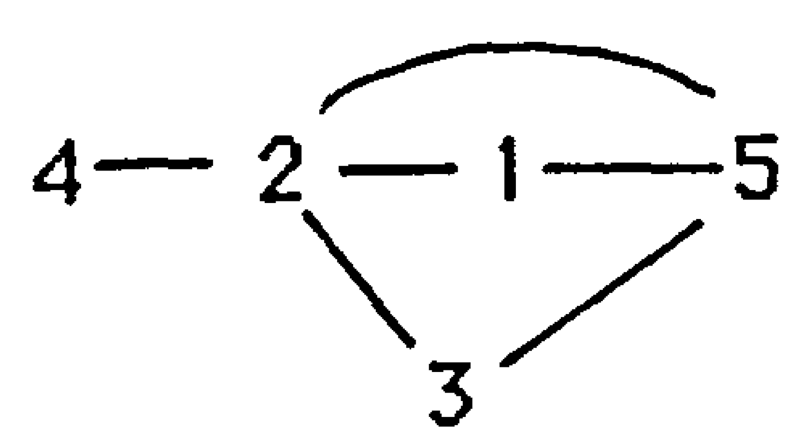
The simplest method of 'clustering' is visual display, or as Anderson (1957) puts it, making devices for helping the eye to aid the mind. The similarity coefficients are grouped into evenly spaced classes arranged in order of magnitude, and each class is given some distinctive degree of shading. A coefficient in a given class is then replaced by its degree of shading. Such a procedure might result in a triangular matrix such as is illustrated in Figure D1a.

By re-ordering, the OTUs which are similarly shaded can be juggled together and clusters can be picked out. The procedure is unwieldy with large numbers of OTUs, but more importantly the arbitrarily chosen class marks destroy information about the actual nodes of the nested hierarchy which is likely to be present.

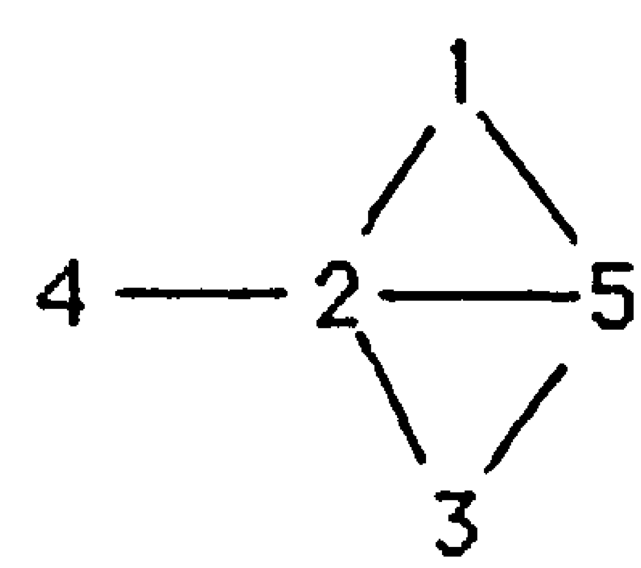
Another possible method is to choose some critical value of the similarity coefficient. Then, when two OTUs have a coefficient which is greater than or equal to this critical value, a linkage between them is scored; when below, no linkage is scored. If we had part of a SIMX thus:

	1					
	2	0.7				
OTUs	3	0.2	0.6			
	4	0.5	0.8	0.3		
	5	0.9	0.7	0.6	0.5	
		1	2	3	4	5
				OTUs		

and we chose the critical value of similarity as 0.6, we would get the following linkages: 2--1, 5--1, 3--2, 4--2, 5--2, 5--3. But we could make this more informative as

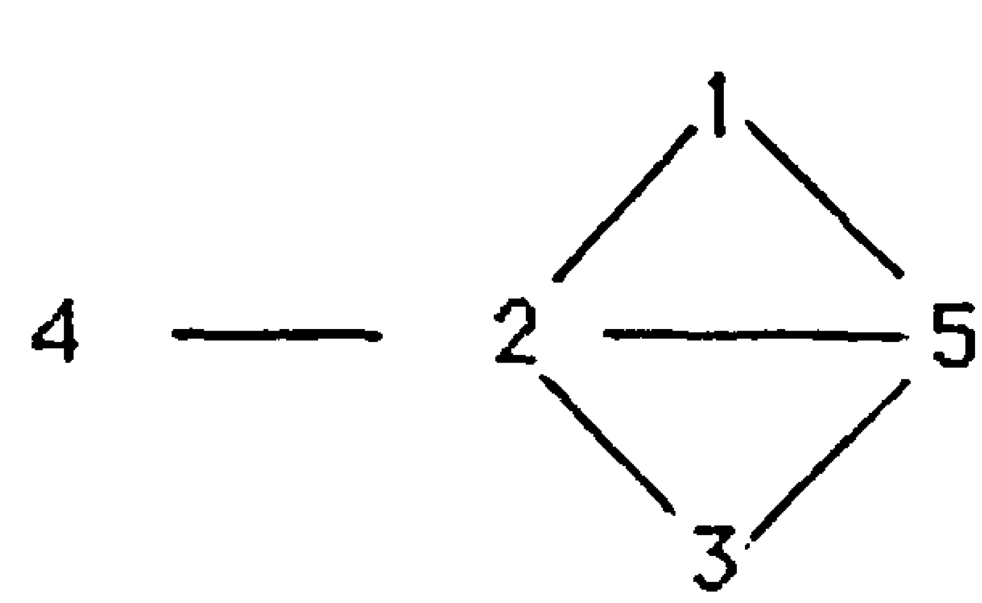


or

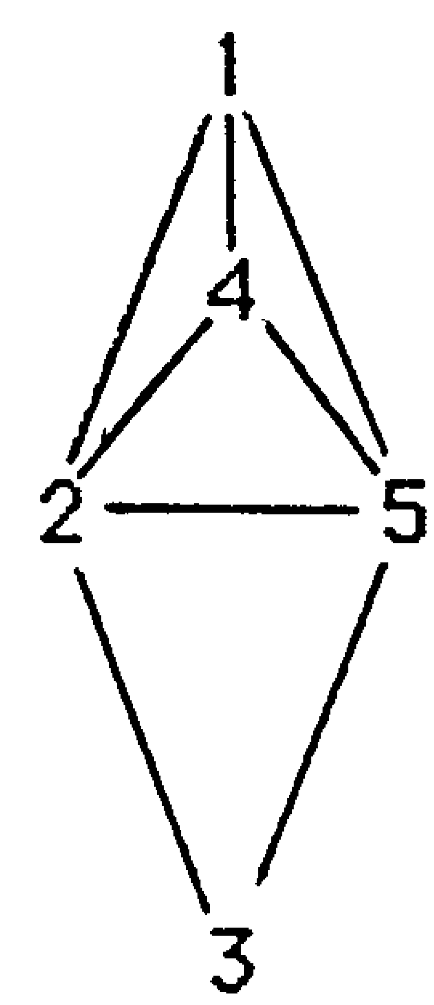


And we might want to interpret this by saying 'OTU 4 is egregious in terms of the properties of the group, but the properties shared by OTUs 2 & 5 are likely to be diagnostic'. Notice how erroneous this interpretive desire is, however, since if we lower the critical value for linkage by 0.1 to 0.5, the linkage list becomes 2--1, 4--1, 5--1, 3--2, 4--2, 5--2, 5--3, 5--4,

and the arrangement



changes to



that is, OTU 4 now ranks above OTU 3 and equal to OTU 1 in terms of 'centrality'.

A dendrogram is a two-dimensional, scalar representation (0-100%, or 0-1) of the successive fusions (or divisions) of a set of OTUs. They have been very widely used in taximetric work, especially biology, but there is considerable disagreement about how useful or reliable they are. Thus Gower (n.d.(a)) says that the assumption that taxonomic distance in hierarchical classifications can be measured is a serious one, and largely unexamined. (But cf. Gower (1971) where it is proposed to compare two classifications of the same data by minimising the sum of squares in order to rotate a set of points each of which represents a distance pair -

one each from separate classifications.)

Carmichael & Sneath (1969) point out that the process of converting an **n** by **N** matrix into an **N** by **N** matrix and then finding connected relatives amongst the **N** involves, at some point in the second of these processes, a form of compression, or, equivalently, a discarding of information, for display. Note that this is different from the loss of information associated with the creation of the **N** by **N** from the **n** by **N**. This latter loss has only to do with the relation between the OTUs and the outside world (i.e. their attributes), not as in the case of displays, with the relations between the OTUs.

They show how this compression for display need not take place between proximity analysis and cluster formation (as is the case in dendrograms), but may take place after cluster formation. They give practical instructions for creating such displays. [N]

[N] (1) Proximities are represented by distance on a suitable scale (0.1 proximity units = 1 inch). (2) Clusters are represented by circles whose D(iameter)s are the same as the d(istance) between the two most distant OTUs in the cluster. (Isolated OTUs are given as single points.) (3) Undistorted d.s are represented by solid straight lines. (4) Stretched d.s are represented by solid straight lines of the proper length joined to their OTUs/clusters by dotted continuations. (5) Squeezed d.s are

represented by solid lines of the proper length, either bent like a 'V', or wavy.
Procedure: (A) Near the middle of the page, draw a circle or point to represent the cluster containing the most central OTU. (Dividing the radii by 100 gives a convenient scale in inches.) (B) Draw the next most central cluster at the given centre to centre d in any direction. (C) A line joining the margins of the circles with centres as points gives the nearest neighbours. (D) Place the next-most central cluster-circle in either of the two possible positions by triangulation from the first two. Connect this one to the first two with lines. (E) Place the cluster closest to any two already drawn by exact triangulation and in the best-fit direction from the others. (F) Repeat till all are located. As each new circle is added, draw lines to indicate its exact and distorted relations with its neighbours.

In spite of their enthusiasm for visuals, Carmichael & Sneath admit that dendrograms produced from one type of linkage will never be sufficient for an interpretation of clusters, since the method is so distortive. They claim that at least two different types of linkage need to be used. Williams & Clifford admit that similarity measures give points of fusion in a hierarchy, but then claim that "although these values are used in the construction of the hierarchy, they are not an integral part of its definition" (1971: 520). Any methods of comparing dendrograms - topologically trees - must therefore avoid being sensitive to conventions of tree representation. Such comparisons must be based on the invariant properties of trees (as given by Berge 1962, Harary 1969). Gower (n.d.(a):

10) concurs: "a dendrogram of given topology can give many different sets of nested classes by permuting the ranking of nodes on different branches, but the number of possibilities is finite", which is, perhaps, some comfort.

Any lingering hope of a definite indication from the dendrogram itself as to what we should believe to be the best structure in, or of, the data is dispelled by Watson et al. (1966: 493) who bluntly tell us that details of the taxonomic hierarchy ought to be decided on the grounds of expediency. But in the botanical case which they are discussing (Umbelliferae and Gramineae) there is a well-known, documented and independent body of practice and belief which supports and promotes that expediency. In our case there is no such experience.

And anyway, there have always been those scholars who see the desire for visual display as at best irrelevant and at worst positively damaging for the progress of taximetric methods. Needham believes that it may be just this simulation of human methods of classification by automatic classification which has inhibited progress - "if a classification is for an algorismic purpose then easy visualisation is irrelevant" (1965b:

114). More generally, Johnson (1970: 210) writes:

> in any but the simplest cases it is no easy matter to elucidate and explain the topological and metrical properties of the infinite variety of definable spaces; there is no reason to assume, as is often done, that simple spatial models are particularly appropriate to taxonomic ordination;

and elsewhere he insists that all we can speak of is relative position in an ordination; that measures of distance may have no meaning at all (when all we have said is that B lies between A and C); that three distinct points define an ordering without any specification of direction in the space.

> If we are working wholly or partly below the level of effective genetic isolation, the tree [dendrogram] is a banyan, anastomosing in its lower levels at least.The sets will then intersect in a complex manner and in this case there is certainly no unique hierarchy . . . In these reticulate cases, which are usually below or about the "species" level in the conventional system, but may include somewhat higher levels if allopolyploidy has occurred, only subjective or arbitrary classification is possible - there is no one "right answer" (Johnson 1970: 231)

****

All grouping methods derive from a complex interaction between (a) the algebraic properties desired of the clusters; (b) the statistical stability of the links between individual OTUs; and (c) the arithmetic value of the starting point. I deal with each of these in turn, although the underlying heuristic and indeterminate facets of each have consequences for the others.

Algebraic properties of clusters.

Before we can choose the best among the paths to be taken through all the variant classificatory procedures so far discussed, we must meditate upon whether we feel there is some suitable mathematical model which fits our intuitions about the data.

All grouping procedures whether they produce groups in accordance with our intuitions in the matter or not, do impose an underlying model upon the data (Needham 1965c; Needham & Jones 1964). The models which are usually used, and whose use is often not justified by their users, are unfortunately very **strong** models, since they assume that the nature of a cluster in an n-dimensional space is simply a function of its 'diameter'. The 'diameter' of a cluster is determined by the arbitrary starting point which must be selected on the similarity coefficient scale (-1 to +1, or 0 to 1). Many workers have tried to remove the diameter problem by progressively coalescing linkages - and thus groups - by working downwards from the pair of OTUs with the highest similarity. This in effect puts them in the position of demanding mutually exclusive groups - each and every OTU is in one and only one group - which are therefore amenable to being represented as trees or dendrograms. Such a method in fact represents an even stronger geometrical treatment than the arbitrary threshold (Sokal & Sneath 1963; Svartvik 1966; Beers & Lockhart 1962).

Needham (1961a,b; 1965c) in work on various forms of information retrieval, found that an algebraic description of the boundary of a group, ignoring the geometrical properties of its shape, gave a much better fit as a model of his data. He also found that the indeterminacy of the clumps (i.e. potential overlap) was productive. A method of finding overlapping groups derived from a graph-theoretical focus on classification has been developed by Jardine and others at the Cambridge

Mathematical Laboratory (Jardine, Jardine & Sibson 1967; Jardine & Sibson 1968a,b). The elaboration of the desirability of overlaps for an improved interpretive framework has been formalised by Cattell (1966; Cattell & Coulter 1966).

Bailey (1982) emphasises the connection between general systems theory and taximetry in terms of the importance of defining groups by the distinctness or otherwise of their boundaries. Conant (1982) illustrates well the sameness of ends but the difference of means between systems theory and taximetry. In nearly decomposable systems the short-run behaviour of each of the component subsystems is approximately independent of the short-run behaviour of other component subsystems, but the short-run behaviour of each of the parts of some sub-system is <u>not</u> independent of the short-run behaviour of other parts of the same sub-system (Simon 1970). Decomposition then reduces to finding sharp changes in an intensity measure of interaction between elements of a system, i.e. a boundary. (Entropies and transmissions are measures of variability and relatedness respectively, but, like most measures in systems theory, they are required to show weak time-invariance.)

Sparck Jones & Needham (1968), in a major test of their clumping program against Cleverdon's document collection (Cleverdon & Keen 1966), show a - for them - characteristic, but all too uncommon, openness about the near impossibility of determining appropriate methods for a large dataset - a considerable number of contrastive experiments will usually be necessary before finding a reasonably small subset of useable methods. They find clumps by minimising the cohesion between a set of objects and its complement in terms of the similarity connections between the two - "a variety of functions are clearly possible" (1968: 92). "The major problem of class-definition [sc. cluster finding] . . . is robustness against shaky data" (1968: 96). Sparck Jones & Jackson (1967) also refer, in their discussion of non-exclusive polythetic classes, to the difficulties presented by large bodies of untidy empirical material. And they say (1967: 35) "it is difficult to find clumps if a substantial proportion of the properties characterise a large number of objects".

It is interesting that Cattell's insistence on a combination of different linkage and starting methods, and his emphasis that two quite different kinds of groups can often be found in the same set of data, are closely related to Needham's views. As far as is ascertainable, their work is independent. Their equivalent methods, perhaps not by chance, arise from the consideration of non-biological data - psychological types, role categories, library classification, semantic classes, and diagnostic patterns in the identification of diseases.

In general therefore, it is extremely important to acknowledge the status of clusters as constructs, and to interpret them in the light of the constraints of the particular model involved. I shall give an abbreviated and simplified account of different notions of "type" (a "type" is a cluster, but one having a different emphasis) and I shall take a path between the set-theoretic exposition of Needham (1961a,b) and the algorithmic one of Cattell (1966).

It is easy to imagine that the multimodal distribution of OTUs in an n-space might be non-normal, but that projected onto the i^{th} dimension it might be normal; it is also obvious that modal density is a relative phenomenon. These supposed underlying numerical properties of modes lead to the notion of quite different but completely compatible _kinds_ of groups or clumps in the same universe of OTUs.

First, Needham defines Kuhn's clumps by:

a set $\underline{C}$ is a clump if (1) $\underline{C} = \{x: x \in N, x > t\}$, and

(2) $\underline{C}^* \neq \{y: y \in M, y > t\}$, where

$\underline{C}^*$ is the complement of $\underline{C}$, and t is some threshold of the similarity coefficient (Needham 1961a; 1965a; Good 1961). (This threshold is the same as Cattell's (1966) 'cutting point'.)

Meanwhile, Cattell & Coulter (1966: 243) define a _homostat_ as "a set of individuals within which the mutual resemblance of all pairs exceeds a certain value, significantly higher than that obtaining between pairs taken at random from the population" (Cattell 1944, 1949; Cattell et al. 1966).

Secondly, Needham defines GR-clumps by:

a set $\underline{S}$ is a clump if,

(1) when $a, b, c, d \in \underline{S}$

and $x, y, z \in \underline{S}^*$,

then, $\underline{S} = \{ a : a \in N, a_b + a_c + a_d > a_x + a_y + a_z \}$,

where a_b is the resemblance between _a_ and _b_,

(2) no member of $\underline{S}^*$ can be added to $\underline{S}$ without contravening (1) (Needham 1961a).

Meanwhile, Cattell & Coulter (1966:262) define a _segregate_ as "a set of individuals which are continuously connected to one another through any intermediate individuals maintaining the stipulated degree of resemblance".

It is clear from these two pairs of definitions that groups of the first kind (Kuhn's clumps, homostats) are not mutually exclusive in terms of element membership, but that those of the second kind (GR-clumps, segregates) are mutually exclusive in terms of element membership. It is strongly stressed by both Needham (1961a,b) and by Cattell (1966) that to arrive at true types it is necessary to proceed from non-exclusive clusters to exclusive clusters.

However various are the definitions of cluster, or however different the purpose of their construction, they all depend upon the series of connections between their constituent members. These connections are often called links, and the process of linkage has certain general properties though rather diverse applications.

Linkage.

There are two general features - whether implicit or explicit - of the linkage process which joins OTUs to form clusters.

First, there must be an arbitrarily chosen value of the similarity coefficient from which to start.

Secondly, there must be some stipulation about the nature of permissible links between a prospective member of a cluster, and the cluster as it stands.

I shall present various applications as abbreviated procedures, with their advantages and disadvantages. (There are minor variants of many of the procedures which I do not propose to discuss.) Sources: Sokal & Sneath (1963); Cattell (1944); Needham (1961a,b); McQuitty (1966, 1967a,b,c, 1968a,b,c) McQuitty & Clark (1968a,b); Lance & Williams (1966a).

(a) Single, arbitrary.

(i) choose a similarity level above which OTUs count as linked, below not linked - suggested 'universally arbitrary values' (Cattell): mean of all positive similarity coefficients, or a series such as +0.2, +0.5, +0.8;

(ii) transform SIMX to a linkage matrix (0, 1) (cf. Fig. D2);

(iii) list for each OTU its relatives in the linkage matrix (LINKX) - this gives single linkage lists (SLL);

(iv) 'ramify', i.e.

(a) take SLL for OTU 1 (which has as relatives, let us say, 2, 7, 15, 16, 17, 23);

(b) take SLL for first relative (i.e. 2) (which has 15, 17, 23);

(c) take SLL for second relative (i.e. 15) (which has 23);

(d) therefore cluster 1 = {1, 2, 15, 23};

(e) take SLL for next OTU not in this cluster (i.e. 3); repeat.

Users: Cattell (1966); Cattell & Coulter (1966).

Disadvantage: Very long-winded, but easy to program - for 200 OTUs at least 60,000 linkage inspections required. Produces overlapping groups

- if this is a disadvantage (see below).

Advantage: And a very considerable one - does not require the re-computation of SIMX after the formation of each cluster. Cattell (1944) gives a 'quick and dirty' version of this single linkage, arbitrary cut-off process in which for 200 OTUs < 40,000 linkage inspections are required. The price paid though, is the loss of small clusters and the absence of appendages or subgroups of 'central' clusters.

Nonetheless, Jardine & Sibson (1968) insist on the following theoretical advantages of the single link method:

(1) It is independent of scale - a multiplier of K simply multiplies the joining levels by K;

(2) the method is independent of the initial labelling (numbering) of objects;

(3) small changes of data produce small changes of result.

(b) Single, highest.

(i) select the pair of OTUs carrying the highest similarity in the whole

SIMX;

(ii) admit more OTUs by lowering the similarity level by minimal - and once established, equal - steps. These newly admitted OTUs are allowed <u>either</u> to join the first pair (now a cluster), <u>or</u> to remain alone as a 'seed' for some new cluster;

(iii) a single link between the OTU under inspection and any established member of a group admits that OTU to that group;

$\underline{S}_j$	OTU groups	
0.99	1,2	
0.98	1,2,3	4,5
0.97	1,2,3	4,5
....		
....		
....		
0.80	1,2,3,4,5	

Users: Sneath 1957b, and others.

Disadvantages: ability of large groups with widely different members at their different edges to coalesce because of a single link. This can be overcome by calculating mean similarities within and between groups, but we have to impose this calculation arbitrarily when we think we are about to lose a group we want.

Advantage: simple to compute (apart from above arbitrary decision); and gives mutually exclusive groups which will make good dendrograms.

<u>(c) complete, highest.</u>

(i),(ii) as above in (b) BUT

(iii) any OTU joining a cluster at S_x must have similarity $>S_x$ with <u>all</u> present members of the cluster;

with provisos:

(1) if OTU_x could join two groups make it join the larger;

(2) if it could join two groups the same size make it join the one which would leave fewest residual OTUs;

(3) if this is indecisive make it join the group with the highest internal mean similarity;

(iv) after all OTUs have joined at S_x recompute similarity matrix from means;

(v) repeat.

Users: Sørensen, Grieg-Smith.

Disadvantages: Computationally complex conditions (especially (iii)(2)). Procedure (iii) so harsh that most clusters would form very suddenly at low similarity levels.

Advantage: Non-arbitrary starting point for recomputation of SIMX.

(d) average, highest.

(i) as above (b) BUT

(ii) admit individuals to the cluster insofar as their addition would not depress the mean group similarity, $\Sigma S_n/n$, by more than an empirically arrived at step (something like between 0.2 and 0.7 (!!)).

(iii) after all OTUs have joined (subject to (ii)), recompute SIMX of all clusters and residual OTUs.

Users: Sokal & Michener (1958); also Rohlf, Sneath.

Disadvantage: Complex recalculation which doesn't always apply.

Advantage: The so-called 'variable group' method - in some cycles no OTUs join, in others many do so. Economical on computer time.

(e) nodal, centroid.

Differs very markedly from all preceding methods since it begins from the single most central member of all OTUs, and not from some **pair** with a given similarity.

(i) from SIMX (whose coefficients are logarithmic) calculate:

(a) $R_i = \Sigma$ of all coefficients > 0 between OTU_i and all other OTUs;

(b) $H_i = (S_{i1}) \times (S_{i2}) \times (S_{i3}), \ldots \times (S_{ij}), \ldots \times (S_{in})$,

where $j \neq i$, $S_{ij} > 0$.

(ii) write OTUs in decreasing rank order of R_i; if $R_x = R_y$, then rank by H_x, H_y values;

(iii) OTU which is highest on R is primary node; OTU which is second highest on R is secondary node, etc.;

(iv) form clusters on primary, secondary nodes etc., by single linkage (see above (b)(iii)). Stop joining to primary node as soon as next OTU to join would be the already established secondary node.

(v) after each addition to any given node a measure of inhomogeneity is calculated for the cluster; if its value drops sharply from the previous value, the last added OTU is removed and the next contender tried instead.

Users: Rogers & Tanimoto (1960); Lockhart & Hartmann (1963) - modified; also Silvestri.

Disadvantage: the computation of inhomogeneity after each addition is lengthy.

Advantage: gives homogeneous clusters, objectively expressive of modal tendency, however heterogeneous or biassed the sample of OTUs was.

Thresholds and stopping rules.

As we saw above, the number of possible pairwise divisions of a sample of n OTUs (in a divisive polythetic hierarchic method) is $2^{n-1}n$, and such a process therefore needs to be stopped before the machine runs out of space-time. It is tempting, as MacNaughton-Smith says (1965: 13), to try to use a significance test on the value of the coefficient as the mechanism for stopping. But this is not appropriate because the coefficient value being considered is not a random one.

Even significance tests on the maximum (or minimum) values of the coefficient will fail because, though the values of the measure may be independent, or even known as to their relations, "the implied population models in numerical taxonomy are too general to permit any such assumption".

The same arguments apply to the setting of thresholds for re-allocation of individuals to groups, or to initiating the fusion of two sets. All such thresholds and stopping rules in taximetry are therefore arbitrary. As a consequence, the simpler they are the better; in the case of thresholds - a numerical value; in the case of stopping rules a simple inequality. (Response thus depends on "the value of some inequality moving in the "right" direction" (Gower n.d.(a): 10).)

Clustering: summary & development.

At the heart of the matter are a difficulty of definition and a difficulty of confidence. 'Cluster' always remains undefined - cf. our initial discussion of Goethe's type. As Ling (1972) remarks, many methods work, but often under unstated and restrictive assumptions. Tests of the significance of clusters are impossible since no statistic is associated with individual clusters. In the face of much that we have read above, Friedman & Rubin's (1967) expectations that there will be innocent consumers of these methods seem strange

> "**N** objects have been gathered. The investigator has no model but suspects there are subcategories of **N**. There is thus no external criterion but the investigator is willing to take an internal one: let the data suggest the subcategories in terms of sets of measures on **N**".

* * *

Clustering proceeds by selecting a value of the similarity coefficient for the starting point and, under various restrictions of the definition of what constitutes a link, adds more and more members until some relatively abrupt change in the numerical relationships between clusters and OTUs takes place. This group is then taken as a cluster. Some form of arithmetical rescaling of the similarities is performed (since remaining

OTUs now have a similarity to a cluster not to its constituent members), and the cycle starts again at a different starting point. Depending upon various definitions and techniques these clusters may be mutually exclusive (i.e. have no members in common) or overlapping. Whatever they are in this respect they are all groups having high within-group similarity, that is they are all Kuhn's or homostat type groups. But, as both Needham and Cattell have pointed out there is often no good reason for believing that the data we are trying to account for are distributed as hyper-footballs, or symmetrical agglomerations of high density in an otherwise empty, or low density, space. Indeed, especially in cultural products (items in Popper's World 3), it must often be the case that distant relatives within some type will have very little in common from an 'etic' point of view. For this reason it will probably not be enough, with these kinds of data, to stop at clusters with high internal similarity, but one will have to go on to process such clusters in search of the continuous connectedness which we should surely be led to expect. This is the progression from Kuhn's clumps (or homostats) to GR-clumps (or segregates). An illustration serves to define two sub-types of homostat and the relation between homostat and segregate (Cattell & Coulter 1966:254-5).

Consider a linkage matrix for 15 people derived from their SIMX on a standard cutting point (Figure D2). By ramifying linkage we can get a reduction to a two-dimensional representation as in Figure D3. Here phenomenal clusters, p-clusters, are those formed at the mere level of observation or linkage - especially that of the single arbitrary type (see above (a), p.492). Sometimes p-clusters overlap and this intersect is called a nuclear or n-cluster. Inspection of Fig. D3 shows that there are different kinds of n-cluster depending upon the complexity of their intersections. It is therefore a structurally informative kind of cluster, though clearly the number and nature of both p-clusters and n-clusters varies with differing cut-off values. But inferences from such structural information must be made with great care since this 'structure' is as likely to be arbitrary, imposed, or procedure-induced as it is to be elicited *by* that procedure *from* the 'phenomena themselves'.

Segregates, on the other hand are sought from continuously connected p-clusters, for which, instead of representing the final list of clusters in a dendrogram, this method gives an overall picture of texture.

Dendrograms have numerous disadvantages. First, they give n-dimensional relationships in 2-dimensional form. Secondly, they are very sensitive to different methods of coding, different similarity coefficients and different clustering methods (Beers et al. 1962). Thirdly, they suppress within-group variation. Textures, by contrast, by emphasising the book-keeping necessities which are attendant on n-dimensional relationships, leave the particular chain of selection and interpretation to the individual investigator, whilst providing an objective basis for replicable results. The 'texture' for a given study is the number of p-clusters of various percentage sizes for different cutting levels; the number and size of n-clusters and their different p-cluster structure; the number and percentage coverage of segregates; the repetitiveness of specific n-clusters, p-clusters and segregates at different cutting levels and so on.

Texture, unlike dendritic treatment, underlines the fact that many typological studies are not necessarily reducible to plane geometry, nor even topology, but answer rather to the continuous often unquantifiable complexity of topography (Cattell & Coulter 1966; Sneath 1967; Thompson 1942).

4. Diagnostic keys and discriminatory features.

It is both useful in its own right, and useful for testing the relevance of chosen methods, to attempt to generalise from the clustered sample to the population at large. This is done by establishing – for all or a relevant subset of the clusters – that particular set of features which best characterises the membership of particular individuals in those particular clusters.

If clusters are of the family-resemblance kind (Wittgenstein 1967), for example GR-clumps or segregates, that is, if they are likely to overlap with each other, then it is likely that there will be no subset of criterial states which is necessary and sufficient for group membership. Beckner (1959:22) writes:

> "A class is ordinarily defined by reference to a set of properties which are both necessary and sufficient (by stipulation) for membership in the class. It is possible however to define a group K in terms of a set G of properties $f_1, f_2, \ldots, f_n$ in a different manner. Suppose we have an aggregation of individuals (we shall not as yet call them a class) such that: (1) Each one possesses a large (but unspecified) number of the properties of G. (2) Each f in G is possessed by large numbers of these individuals and (3) No f in G is possessed by every individual in the aggregate. By the terms of (3), no f is necessary for membership in this aggregate; and nothing has been said either to warrant or rule out

the possibility that some f in G is sufficient for membership in the aggregate."

This absence of a constant feature in a cluster complicates the construction of keys, but does not make it impossible.

If Beckner's definition holds in full for the clusters found, they are said to be 'fully polythetic' (thetos – an arrangement of , in this case, criteria / individuals with respect to individuals / criteria). If only the first two conditions hold for a particular cluster it will be partially polythetic. Depending upon the degree of polythetism in any set of clusters, there are two separate routes to a diagnostic key. Either a multiple version of Fisher's discriminant function can be used when the clusters are fully polythetic (i.e. all members of a cluster have no single feature in common); or when the clusters are partially polythetic, they can be converted into monothetic ones, thus giving discriminatory features directly (Lockhart & Hartman 1963).[N]

[N] Fisher's discriminant function has the general form $F = L_1x_1 + L_2x_2 + L_3x_3 \ldots + L_nx_n$, where $x_1, x_2, \ldots$ are the criteria and $L_1, L_2, \ldots$ are constant multipliers determined such that OTUs allocated to one cluster have their discontinuity from OTUs of another cluster maximised.

According to Sokal, an important limitation of the discriminant function is that it requires an _a priori_ separation of the individuals into recognisable groups (a known population). Furthermore, it takes account only of those attributes whose means differ in the population, i.e. have overlapping frequency distributions. Gower (1969) claims that discriminant analysis is most useful when the populations overlap - since it tries to assign individuals in terms of minimising a misclassification metric. "When the distribution of the [attribute] values does not fall into any of the classes, then practical techniques of discriminant analysis do not exist, although the mathematical theory is known" (1969: 371). Hall (1968) proposes a method for finding weightings for quantitative attributes from a discriminant function, but there is, of course the Never Evitable Catch. His method depends upon fusions of sets which must always have equal numbers of members !

The general desideratum of a method of diagnosis is a definition of groups such that criteria for the allocation of contingent members are supplied, together with an estimate of the probability of correct allocation.

Compared with the amount of published material on the practices and problems associated with coding, similarity estimation and clustering, very little work has been done on the scientific construction of diagnostic keys. Sokal & Sneath "only point out a few of the possible lines of development", stating that in general it has been usual to return to "the original tables of data in order to find those characters which are the most useful for rapid identification" (1963:275). Cattell et al. (1966:324) are also little more than programmatic, incorporating a discriminant function to alternate with the clustering process so that, in some sense, the clusters are the diagnostics;

> "we propose ... the iterative multiple discriminant function ... technique in which [the clustering process] is followed by the multiple discriminant function, and one then reverts to the [clustering] procedure after weighting the dimensions [sc. criteria] in the [similarity] calculation with the weights found in the multiple discriminant function, and so on to some iterative convergence. At present this would be laborious ..."

Beers & Lockhart (1962) propose a less laborious method which is computationally much faster. Given that the standardised feature matrix can be retained, P the proportion of positive criterial states within any given group, is simply a question of look-up in that feature matrix.

P is given by: $P_i = n_i / N_x$,

where n_i is the number of positive responses of criterion **i** in group **x**, and N_x is the number of OTU members in group **x**.

A catalogue of P-values for the group would be the diagnostic description of that group. Values of **P** near 0 or 1 (minimum or maximum) would be key (negative or positive) diagnostics.

Carvell & Svartvik (1969) propose three general methods of key construction which may be used singly or in combination. All three methods assume polythetic clusters, but adopt different starting points and produce keys with slightly differing members, or ordering of members. They are:

(1) "polythetic key", directly from output;

(2) approximate "monothetic", indirectly from intergroup analysis;

(3) concept-formation key, indirectly by feature selection.

All three methods operate upon the basis of critical phenon levels - that is, places in a dendrogram where major changes of cluster structure take place.

Thus we might have the simple dendrogram pictured in Figure D4, where the ordinate is % similarity (i.e. similarity coefficient x 100). Here we have three 80% groups: (1,7), (4,3,2,8,6,11,13) and (5,12,9,15,14,10); which, at the 85% level become seven groups: (1,7), (4,3), (2,8,6,11,13), (5,12,9,15), (14), and (10). **But** these seven groups would also 'come out' at the 83% level, therefore 83% similarity is a critical level for these groups of OTUs. It is clear, both logically and from personal experience that at just these points where major clustering changes occur, changes in the co-occurrence of criterial features also manifest themselves.

Carvell & Svartvik (1969) impose restrictions on the nature of groups which can be examined at any critical level. First, they must have more than a certain minimum of members (Carvell & Svartvik suggest 5 - their total sample was 146 OTUs); secondly, they must be subgroups of the next highest retained group (perhaps 2/3 of the number of members of the next highest group).

Given that we have a cluster satisfying these requirements, method (1) - polythetic key - proceeds by applying two independent measures to the criterial responses of the group members.

Polythetic groupings.

The measures are statistical tests which attempt to ensure that criteria showing a high within-group relevance will also have relevance for the (isomorphic) classes in the unsampled population. The tests used were χ^2 and the diagnostic value, D_V.

Only very high levels of χ^2 were accepted since in fact most criteria had a χ^2 value of above 95% in their experiment. Carvell & Svartvik admit that χ^2 is not a particularly suitable test. There are three reasons for this:

(a) the k samples (criteria) are not independent in terms of the clusters formed: independence of k samples is the primary indication of the relevance of χ^2;

(b) the use of 'particularly' high values of χ^2, as representing diagnostic criteria, is an arbitrary choice of a subset of values from a range of confidence-limits which is continuous;

(c) its use is circular. It establishes a degree of association between individuals in the group, but it also gives both the significance of feature-group association, and group reality or stability.

The χ^2 test of the diagnostic power of criteria is therefore supplemented by a measure of diagnostic value, D_V. This, in information theoretic terms, predicts the likelihood that a particular criterial response will decrease the amount of uncertainty attendant upon allocating a new individual to a given class,

$D_V = (\log [\,^{\circ}{}_{I=K}.\{f_1, f_2, \ldots, f_n\}]) - \log_n / (\log [\,^{\circ}{}_{I=K}.\{f_1, f_2, \ldots, f_n\}])$,

where $(\log [\,^{\circ}{}_{I=K}.\{f_1, f_2, \ldots, f_n\}])$ is the "overall uncertainty" inherent in any particular established group; º means 'permute'; $_{I=K}$ 'keeping Inventories (k) constant'.

The idea arises from the following considerations: in a given group we know the proportion of members having any particular criterial response, and we know the number of members. The uncertainty, or ignorance, inherent in the group is then the log. of the number of ways the responses

could be distributed amongst these OTUs whilst holding the numbers (of responses, and of OTUs) constant. The diagnostic uncertainty of a criterion is $\log_n$, where n is the number of ways OTUs can be put into groups, subject to the knowledge that these OTUs carry known responses to the criterion in question.

Approximate monothetic.

Taking all relevant groups at a critical phenon level of the dendrogram, a similarity matrix is computed between those groups, and the intergroup similarities converted into a linkage diagram (the first union). Groups from the first union (G_1) are thus groups (G_1) of groups (G_0). A second union, if possible at all, is made along the same lines, or by inspection.

For the groups of whichever union is terminal (G_1, G_2, G_3, . .) construct a table in which column 1 gives the mean internal similarity of the groups; column 2 gives the inter-group similarity of groups and their complements; column 3 gives mean internal similarity of group complements. Columns 1 and 3 should be high, column 2 low. The terminal group which has mutually highest values in columns 1 and 3, and lowest in column 2, is the best partition and should, in conjunction with its complement, give the primary diagnostic features for the whole study. Extraction and ordering of diagnostics proceeds from best to worst natural partitionings of terminal groups, iteratively, back to the original clusters.

Concept formation.

This proceeds substantially as (2) above, but is preceded by 'hand' selection of the main features of each primary cluster. These features are then used to define similarity coefficients which establish intercluster relationships.

* * * * *

Diagnostic methods of a rather different form are proposed by Lockhart & Hartman (1963). In effect, by forming the groups monothetically, the amount of computer time which they spend clustering is radically reduced, and the method of linkage gives clusters which embody the diagnostic key.

A nucleus pair is a pair of OTUs in the SIMX having very high relative similarity (i.e. few criterial differences). Instead of the usual methods of linkage however, the following transformation is effected. One makes the assumption that properties not shared by the nucleus pair are inessential, and they are discarded. If n_{dn} is the number of properties shared by this dyadic nucleus, then $n - n_{dn}$ properties are discarded. Next a search is made for a third OTU which has fewest differences from n_{dn}. Properties not shared by all three are now discarded (i.e. $n_{dn} - n_{3n}$). A fourth OTU is now sought differing least from n_{3n}. As each new member is added, a printout of its number and the cumulative difference, d_c, is required. For the fourth member added above, the cumulative difference is

$$d_{c4} = (n - n_{dn}) + (n_{dn} - n_{3n}) + (n_{3n} - n_{4n}) .$$

There are several useful advantages in this method, in addition to its saving of time:

(1) all OTUs are similar in the same respects at any given d_c level;

(2) the polythetic groupings and the monothetic groupings of the same OTUs were similar (in this experiment: bacteria);

(3) when polythetic groupings are computed from the semi-metric distance coefficient -

$D_{RT} = \log_2 .(n_s / n_s + n_d)$, where n_s includes negative matches (see above) -

the relationship between polythetic and monothetic clusterings is maximised. (That is, plotting D_{RT} against d_c gives a linear regression (cf. Lockhart & Hartman 1963; Rogers & Tanimoto 1960).)

(4) the computation can be even further shortened by selecting an 'artificial nucleus pair', not the highest (true) nucleus pair, and (sometimes) getting the same groups.

For a given BDRM there is no unique diagnostic key, but some are more useful than others. But all keys depend on a BDRM which is as full of entries as possible; and the values of an attribute to be used in a key must be mutually exclusive, and therefore, by inference, quantitative attributes will be difficult to use in a key (even if they have been dichotomised) (Gower 1969). Pankhurst (1970) provides us with an algorithm which will heuristically select the optimal key of all keys in a tree searching process. The algorithm generates another algorithm, and the algorithm which is generated is the key itself, which may be searched and will accept grafted branches. Ideally the subgroups at each point of division are (a) two in number, (b) equal in size.

Diagnostic keys: summary.

The selection of features from already constructed groups which will correctly allocate newly sampled individuals to their proper group proceeds in one of three general ways:

(1) selecting by hand;

(2) applying certain statistical/numerical tests to the co-occurring criterial responses of polythetic groups;

(3) either (a) converting polythetic groups to monothetic ones;

or (b) computing monothetic groups from the SIMX in the first place.

5. The adequacy of taximetry.

We have already seen that there are various difficulties and indeterminacies at many points in the business of choosing a suitable set of techniques for classifying one's data. Here I merely draw them together under eight headings.

(a) Taximetric adequacy: comparability and repeatability.

Sometimes we may wish to compare two or more classifications. Reasons for wanting to do this are many and varied, but we may divide them into two groups for our present purpose: (a) comparison for the sake of examining the effects and effectiveness of different methods; (b) comparison for the sake of the results of classifications. Of these I shall concentrate mainly on the first (though for an interesting unsolved problem of the second type, see Pellowe (1973)).

Williams & Clifford (1971) suggest that when we compare classifications, we should draw a distinction between those which differ conventionally, and those which differ intrinsically. For instance the values of a comparison function (CF) which enable the construction of a dendrogram in two dimensions are irrelevant to the definition of that dendrogram (as are the two dimensions) - the only relevant matter is the ordered string of successive fusions (or divisions), that is, the topology of the tree. Bobisud & Bobisud (1972) make much the same point, but argue that the methods of Williams & Clifford (1971) limit the possibility of comparison to dichotomous classifications; using Harary's (1969) graph theory they generalise the method.

Morgan (1973), on the other hand, shows how the use of two methods (multi-dimensional scaling and single link analysis) on the same SIMX can produce mutually complementary and informative results.

Sibson suggests a way of rejecting noise from a classification by using secondary comparisons. His logic is: if _a_ is highly similar to _b_, then the cohort of _a_'s similarities to others ought to be similar to the cohort of _b_'s similarities to others. Thus comparing _patterns_ of similarity values ought to yield a CF similar to the original CF, but with the wrinkles smoothed out.

Gower provides a test, which may well be too powerful, by suggesting (1969) that if a method can lay any claim to generality, then similar classifications should be obtained when different sets of characters are used from the same OTUs. Sparck Jones and Jackson (1967: 35) are much gloomier about this:

> one cannot even expect a flexible general-purpose classification program to give good results with bad data, or even good results first time with good data, unless any major characteristics of the data which will influence the classification have been allowed for (say by weighted similarities). These characteristics are often not very obvious.

In an important and painstaking experiment, Moss (1971) quantitatively compared the intuitive and numerical classifications of the same sets of OTUs by groups of workers in two different institutions who had received comparable training in classificatory and taximetric technique. He assumed that the pooled intuitive (I), and numerical (N), classifications of 11 University of Pennsylvania workers and 12 University of California workers was the closest approach to the "truth" that could be achieved with either the intuitive or the numerical method of classifying. Both methods were undertaken by both groups on a sample of 7 OTUs and on a sample of 12 OTUs (which included the 7 of the first set).

Moss found:

(1) that the overall pattern of skewness and kurtosis of the similarity coefficients didn't change markedly for change in number of OTUs;

(2) that in I classifications close similarity is accentuated by selection of too low a distance, and dissimilarity by too high a distance; [N]

[N] Moss (1971: 325) notes "it seems reasonable that the mind would tend to exaggerate close- and de-emphasise distant- relative relationships in an attempt to recognise subsets within a set of OTUs . . . the major groups were more distinctly set off by the intuitive method and the [similarity values] spread over a greater range than that shown by numerical values". We may suspect that this tendency would become more marked as the number of OTUs rose - leading, no doubt, to substantial between-group vagueness (D,C8).

---

(3) that as the number of OTUs increases, the correlation increases - in both I and N methods - between the mean similaritites of distant relatives;

(4) that the I method is better than the N method for a small number of OTUs;

(5) that the I method is less variable than the N method for the highest mutual similarities, (N is very variable for within-group similarities);

(6) that N is less variable than I for between-group similarities.

These results speak for themselves.

The experiment failed to demonstrate the existence of sub-groups of workers. Moss is disquieted by this;

> "if the training received by students can be at all related to their ability to infer similarities, then one would expect that graduate students would do at least a slightly better job than undergraduates .. Either [the] assessment of similarities is independent of instruction, or [the] instruction was not entirely adequate" (1971: 326).

Given the pervasive use of classificatory activity in daily life one might imagine that the former was more likely.

And this is implicit in Moss' own rejection of the notion that taxonomies of biological OTUs and non-biological OTUs must be different in kind: a notion in whose favour I have never read a satisfactory argument; but against which the internal methodologies of folk taxonomies speak eloquently (cf. Berlin et al. 1974; Hunn 1977) (D,C9).

In comparing any two classifications it is important to separate, but also to take account of, the level of the relationship and the shape of the relationship between two groups, or dendrograms, or whatever is being compared (Fleiss & Zubin 1969). Williams & Clifford (1971) use the terms constitution and configuration in an analogical way; and say that when we want to compare classifications we must find some measure which is roughly additive over both these features.

In ecological studies it is sometimes useful to compare the results of two classifications - species as to their abundance and co-locations and habitats as to their niches and populations. In order to do this it is necessary to consider the results of normal and inverse association analysis together. (Given such a treatment, linguistic varieties and social situation could then be examined as truly interacting (Pellowe 1973).) For the association between species and habitats, Williams & Lambert (1959,1960,1961,1962) do this by hand, but the method in computation would be to use a two-way table of **superimposed** Q-technique and R-technique results. But, needless to say, there is a Catch:

(a) if the table is standardised then the axes are at maximum variance, but the analyses (R- & Q-) become non-equivalent;

(b) if the table is not standardised then the axes are not at maximum variance, and though the analyses are equivalent, the purpose of superimposing the original analyses is lost.

(b) Taximetric adequacy: resemblance.

Sibson warns that it is misleading to operate on CF values (similarity/dissimilarity) as though they were distances. Gower (in Sibson 1972) says that it is the properties of discovered classes which are of primary interest, and distances between classes, though sometimes fruitful as a way of looking at a problem, do not contain all the relevant information.

Shepard (1963) claims that non-metric information "contains" metric information in such a dilute form that we can't see it. If you "squeeze" the non-metric information you see the metric in it - reduction is essential.

And Gower (comment on/in Sibson 1972: 342) says that non-metric methods may not be worth the extra computation.

The wholly undefinable, abstract nature of resemblance is insisted upon by Johnson (1970). Because of the infinity of attributes, and because of the infinity of ways of establishing resemblance, there can be no true or natural overall similarity between objects. Restrictions of an entirely subjective nature may lead to a set of measures, but we should never deceive ourselves into thinking that such measures are isomorphic with resemblance as a phenomenon of gestalt.

(c) Taximetric adequacy: stability of classifications.

In comparing classifications we are, in one sense, comparing the stability of the algorithms which produced those classifications (Bobisud & Bobisud 1972). And we might imagine that an algorithm was stable if the addition of new objects did not result in any change in the classification of the original set of objects. This seems impossibly powerful to me. And cf. Gower's surprisingly optimistic remark (1969: 363) "similar classifications should be obtained when different sets of characters are used" - against which Minkoff's (1965) work militates. Bobisud & Bobisud go on, not surprisingly in my view, to say that a stable algorithm probably does not exist, and that no way is known for determining abstractly how nearly stable an algorithm might be. That determination can only be done empirically.

Jackson & White (1971) point out that there is an important connection between stability and continuity. A classification is produced by a transformation of the data; small changes (of data) will be represented as small changes (in the classification) only if the transformation is continuous. Stability is then not only a property of the algorithm, but also of the population and its properties. Johnson (1970: 223) in characteristic fashion points out that we can have little idea what stability itself means since we can get "considerably divergent classifications of the same material from the same data". And this problem is compounded by Williams & Clifford (1971) who sense that stability and classificatory power (in a method) are incompatible. There seems to me to be a considerable possibility that they are **inverses** of each other (D,C10).

(d) Taximetric adequacy: optimality.

Clearly there should be a certain amount of overlap between the idea of stability in a classification and the usual yearning for its optimality. The problem, as usual, is finding an external criterion of "bestness". Gower (n.d.(a): 3) asserts that the best classification is the one that predicts most characteristics correctly. MacNaughton-Smith (1965) says that the internal criterion for a good taxonomy is stability in the face of fresh data (attributes or OTUs), and that the external criterion is that the taxonomy lead to hypotheses which withstand testing on competely fresh data.

An optimal classification, according to Warburton (1967) will maximise the probability that statements known to be true of two organisms are true of all members of the smallest taxon to which they both belong. It should not be impossible to develop tests for this property to objectively decide which of several rival classifications is best.

MacNaughton-Smith having said that there is no such thing as a general (i.e. natural) classification, writes "within very broad limits it is meaningless to suggest that any particular method of proceeding is "invalid" " (1965: 7). Watson et al. (1966) conclude that

> "the most significant fact to emerge . . . is that anybody, given access to comparative data and any one of our crude processing methods, might well have produced a taxonomic scheme representing a slight improvement on [the best 19th century intuitive classification]",

but then, anybody might just as well **not** have !

(e) Taximetric adequacy: inference.

There is an increasing amount of debate amongst taxonomists about what classification and classifications should or should not be used for. We have already seen that Moss does not believe that biological and non-biological classifications are different in kind. Watson et al. on the other hand think it dangerous to draw parallels between ecological communities and taxonomic groups, since the latter have evolution as an underpinning of their predictive, natural, real force (1966)[N].

[N] But 'evolution' itself is surely not a sufficiently secure notion to be able to remove the citation marks from 'natural' and 'real force'. (Cf. The conspiracy of silence, as far as the text-books and the people are concerned, which makes it seem as if geneticists believe in Darwinian 'evolution'.)

Friedman and Rubin urge caution when there is no external test of the structure in the data, for instance a classification from some other source,

> "if there was little prior knowledge about the group structure of a set of data, we would analyse this data with several different criteria [sc. measures, stopping rules, etc.] before coming to any strong conclusions" (1967: 1177).

An example of inference which goes well outside the warning of

Friedman & Rubin is Sibson's (1972: 317) speculation about evolutionary divergence:

> we know virtually nothing about rates of increase of dissimilarity over time, but what we may sometimes be able to assume is that if y diverged from x before z did then the dissimilarity between x and y is at least as large as that between x and z, and conversely; in other words ideally there ought to be local order equivalence between present day dissimilarity and time since divergence.

(f) Taximetric adequacy: models and metaphors.

Classifications are reflections of the classifier's interest or purpose. The importance of fully understanding one's interest and hence one's purpose is emphasised by Jardine & Sibson (1971b) who conclude that controversy concerning criteria of adequacy for methods of automatic classification can be partially resolved by considering the purposes for which scientists classify. Usually, the aim is to simplify data in ways which may suggest or confirm hypotheses; but occasionally a worker will want to impose external constraints such as wanting homogeneous clusters, or a particular size or number of clusters. These constraints are incompatible with the sort of adequacy associated with data simplification.

If a scientist wishes both to obtain homogeneous clusters and to

> obtain a simplified representation he should use a non-hierarchic cluster method. If he definitely requires a hierarchic system of clusters he must make a choice between adequacy of representation and homogeneity of clusters by considering carefully the purpose for which he is classifying (1971b: 406).

Objecting to Gilmour's notion, that the value of a classification can be assessed by means of considering the range of its purposes, Johnson (1970: 213) notes that other animals classify, and "so do some machines; certainly any highly intelligent being would do so (though . . . some might be less compulsive classifiers than ourselves)". All classifiers classify for the sake of some interest, but one person's interest may not be another's and thus 'range of purposes' must remain undefined. Furthermore it might be asserted that the interest determines the classification. Johnson again (1970: 217)

> when we select (and we do select) the acceptable attributes to be drawn upon in making a "general" taxonomic classification, we also unconsciously choose a particular set of attributes specially reflecting a "factor" which is strongly linked to our (selected) interests.

The notion of interest can be examined better perhaps by considering the confusion we feel at its apparent absence. Thus Johnson (1970: 216) invites us to attempt a "general" classification of the object set:

{Mao Tse-Tung, a peanut, the Sphinx, an electron,

a litre of alcohol, the star Achernar};

and he predicts that "we shall find it very unsatisfying because we cannot define the field of our interest".

Of course we may think of classifications in other ways than as the basis for recognition keys or the source of testable hypotheses. They may be considered as teaching or learning devices - in the form of an organisation of masses of otherwise unconnected facts, in which classes of relations, as well as of things, are represented (Johnson 1970). (Consider Mendeleev's Periodic Table.) Watson et al. (1966) see a good classification as a convenient framework for storing observations which will, themselves, often be predictive. In a multidimensional scaling of various confusion data for Morse code signals, Shepard (1963) suggests that the reductive nature of the process - turning non-metric information into metric by trading off degree of dimensionality against monotonicity - reflects some information processing characteristics of man. Johnson (1970) concurs, being of the opinion that our ordinary intuitive processes begin to lose efficiency with problems of this magnitude [200 objects & 200 attributes of various types].

> Perhaps if our mental apparatus were differently organised multidimensional ordination would meet most of our requirements for the organisation, retrieval and comparison of information. This is a question for those studying the design of logical and quasi-mental machines (1970: 212).
> [Caianiello (1961) addresses this problem in part.]

In spite of our compulsion to classify in daily life, and our addiction to analogy and extrapolation, the space of the classifications which we make may differ radically from physical space. That is, one of the most fruitful characteristics of analogy is precisely the ways in which it challenges the notion of an identity between itself and that which is analogised. Classification space may differ from physical space in one or more of the following respects: it may

(a) be finite,

(b) have more than three dimensions,

(c) be non-Euclidian,

(d) be non-metric, (or semi-metric, or quasi-metric),

(e) be discontinuous.

The N-dimensions may not be an extension of Pythagoras, but may be in the geometry of Riemann (in which at any point on a line there is **no** line perpendicular), or in the geometry of Lobachevski (in which at any point on a line there are **two** different perpendiculars). The shortest distance – like the rook in chess – may be round a corner. No one or other of these notions of space/distance can a priori be said to be better or more real than another.

(g) Taximetric adequacy: inexactitude.

Science is actually more interested in qualitative patterns than quantity. Group theory and topology are more relevant branches of mathematics than statistics; isomorphism and the matching of sets of relations are of more significance to theory development than frequency distributions.

> We should not be bewitched by number, in particular the continuum . . . we shall probably never be able to express or apprehend complex situations in their precise quantitative detail (Johnson 1970: 229).

Sibson (1972) echoes this by pointing out that the stochastic model which underlies classical statistics is missing altogether in classificatory studies or plays only a very subordinate role. Whitehead sees classification as mediating between perceptual processes and mathematical abstraction:

> Classification is a half-way house between the immediate concreteness of the individual thing and the complete abstraction of mathematical notions. The species take account of the specific character, and the genera of the generic character. But in the procedure of relating mathematical notions to the facts of nature, by counting, by measurement, and by geometrical relations, and by types of order, the rational contemplation is lifted from the incomplete classifications involved in definite species and genera to the complete abstractions of mathematics. Classification is necessary. But unless you can progress from classification to mathematics, your reasoning will not take you very far (1929: 37).

Johnson makes the same point the other way up. He says that there is no hope of our being able to extend precise quantitative mathematics, such as tensor analysis, to describe the situations encountered in taxonomy. This is because the subject matter of taxonomy is so complex. "Mathematics is hard, paradoxically, because its subject matter is simple" (1970: 229). He believes that in the most difficult applications qualitative assessment - which doesn't exclude mathematical concepts - may be a surer winner than quantity. Many other tools and techniques, new and old, are at the taxonomist's disposal

> including the subtlety of the human intellect and its power of perceiving Gestalt and of bringing information and theoretical reasoning of all kinds to bear on a problem (1970: 235).

Systematics is a stimulating and rewarding, though inherently inexact branch of science.

(h) Taximetric adequacy: subjectivity.

The results of computer generated attempts at classification are useful as a heuristic for suggesting a classification, since there is nothing objective about any classification produced in this fashion. Cormack (1971: 405) asserts that any attempt to generate clusters which are more homogeneous than those produced by single link analysis leads to arbitrariness in the cluster method. MacNaughton-Smith (1965) admits that we have little experience of the predictive efficiency of all these techniques under different conditions "and the choice of a method still depends to some extent on hunch" [emphasis added]. Malone (1975, & p.c) points out that many of the techniques in his programs operate without a model and make no use of prior knowledge, even if there is any.

> They have many "degrees of freedom" and a great capability of "capitalising on chance" which implies relatively low reproducibility of exact results ... Little statistical theory exists to directly assist in interpretation of results ... [and so one relies on] "gut feelings" as to the plausibility of results (1975: 5-6).

Fleiss & Zubin (1969) give three arguments against using cluster analysis to find underlying types:

(a) there is no mathematical model whatever;

(b) there is rarely a statement of exactly what one is looking for;

(c) the distribution of the sampled population is not known and is unlikely to be normal.

And Friedman & Rubin (1967: 1162 & 1166) show the total absence of straps for booting with (the Catch of Catches) when they say :

> if we knew the groups we could define an appropriate distance measure; if we knew the appropriate distance measure we could manage to find the groups... Until we know the groups we cannot select a relevant subset of variables; until we have selected the variables we cannot inspect their effect in group formation.

It is finally unsurprising that Lincoln Constance could say "Taxonomy in its widest sense, which goes far beyond the formulations of classifications, remains "an unending synthesis" " (Johnson 1970: 235).

6. Deciding on a taximetric strategy for linguistic variation.

It is clear from all of the preceding that the set of strategies which is best suited to the data of linguistic variation is not immediately obvious. This perception is, even given our present purpose, as revealing of our underlying ignorance about the appropriacy of any given modelling of linguistic variation as it is of the difficulty of understanding the detailed consequences of a set of particular methodological choices within any particular classificatory scheme. Any given classificatory method will tend to **elicit** from a given data set its own kind of classes. These classes will have those geometrical and algebraic properties which are predictable from the details of the method, rather than those properties which are inherent in the data. And these kinds of classes will be elicited by that method regardless of the likelihood of their being present in that data set. This being the case, we are forced to ask ourselves in some detail what we **expect** the lumpiness or clumpiness of linguistic variation to look like. Having certain definable expectations in the matter, we can select or design an overall classificatory method which optimises the elicitation of **that** kind of clumpiness from the data. And finding our expectations by and large reflected in the clumps we get, we can engage in the satisfying busy-ness of to-ing and fro-ing from results to theory, claiming that the former vindicate the latter and using disparities between the two - which we would very probably insist on calling minor - to celebrate, elaborate on, or expand the latter.

Regrettably things are not so far in hand that even this prospect appears much more than a pleasing dream, since it proves difficult to decide **what** sort of clumps we expect. In other words, we don't know what sorts of distributions are involved in whole-system and sub-system variation, nor within and between differing kinds of groups of speakers in the population. At least we don't know about these things in any manner which might lead to the selection of any definite combination of:

(a) a format for the expression of the variates (data coding),

(b) a measure of similarity,

(c) an initiation method,

(d) an allocation method,

(e) a stopping rule,

(f) a reallocation method, and

(g) a method of deriving a diagnostic key.

And even if we could delicately and appropriately make such decisions, then our 'results' would only constitute a complex and, by definition, grossly inaccurate and unwieldy fiction produced by the left hemisphere (sc. for right handed persons) in its attempt to mimic so wearyingly and egoically that which is accomplished wordlessly, unconsciously, holistically, almost instantaneously, incomprehensibly, and certainly unmodellably, by the right hemisphere.

Nonetheless the Way, which **may** lead us out of this endless thicket of uncertainty, must be one which depends upon feeling the appropriacy of the **obverse** of Labov's (1972) assertion that the highest goal is to be right. This obverse grants that humility will always admit more ways of seeing things than power ever can (Tao Te Ching, passim (Lin 1977)).

ALL MISSING PAGES ARE BLANK

IN

ORIGINAL

Appendix E

Establishing a social space

The banality of the coding frame for 'social information' which follows is somewhat mitigated by the arguments presented in Chapter 3 above.

It represents a perhaps not very imaginative attempt to provide a social space in which people would be able to place themselves, or even to recognise themselves as being closer to or more distant from other point-persons within the embrace of its dimensions.

Parallel proposals are to be found in Lewin (1936) and Herbst (1973). Lewin proposed a life-space concept which, however, was never made operational, whilst Herbst, in grappling with the empirical problems suggested by Lewin's notion, found the need to transform the life-space concept into his behaviour-system concept. In either case, the goal was to represent the situation that 'exists' for a person as s/he perceives it and/or as s/he believes it to be.

Most sociological and psychological theories take their basic assumptions from sciences which deal with non-organic phenomena of cosmic or atomic proportions. Far too little attention has been paid to the likelihood that "a theory of [e.g.] physics cannot serve as a general model for scientific method" (Herbst 1973: ix). The notion that one can find a set of laws which apply everywhere in the universe, which are based on the same set of measurement scales, and which are time-invariant is clearly a notion which did not evolve from studies of the behaviour of human individuals or groups. No anthropologist, for example would imagine that such a notion had any use in his professional life (cf. e.g. Bateson 1979a).

And even in physics, as we now know (DeWitt & Graham 1973; Dirac 1982), the measurement scales change in different circumstances, and the 'laws' are variant depending upon the size of the unit, and upon the motive for, and time of, its observation. Encouraged, Herbst notes:

> properties of this kind, instead of being unusual, turn out to be essential characteristics of human behaviour . . . Every person and every group has the characteristics of a behavioural universe which evolves its own laws and measurement scales; . . . [and] even for the same person or group the principles of behaviour change over time (1973: xi, emphasis added).

But after examining many cases in support of his expectation that "it would be possible on the basis of general system theoretical concepts to formulate a theory of behaviour applicable to all persons and groups" (1973: xi), Herbst abandoned his attempts to implement this possibility, since

> if the data for [some group] are summed together, a stable pattern of relationship between behaviour variables emerges. However, the results obtained in this way bear no relation to the pattern of relationships found for any individual [of that group]. This leaves the problem of how data obtained by population-sample studies can be interpreted [since such data represent] properties of populations but not necessarily . . . characteristics of individual human behaviour . . . The only alternative would be to construct a different behaviour theory for every human being (1973: xi, emphasis added).

Clearly, for all of these dimensions, there is no natural cut-off point at which an increasing delicacy of gathered information becomes irrelevant to the construction of a Social Space which will appropriately and recognisably (too strict a condition ?) locate any individual in the population. Equally clearly, the space suggested here should, in addition to the dimensions which it does try to incorporate, add further sub-spaces; perhaps of personality traits, perhaps of

preferences in terms of cultural and artistic productions, perhaps of aspects of political and ethical thinking, perhaps of general intellectual ideology, perhaps of preferred companions, even, perhaps, of preferred foods, and so on, and so on. And such concerns as the ones I have incorporated are certainly not restricted - either as to their fascination or as to their importance - to those who might characterise them in such a fashion as I have given.
Without some further components of these sorts, how could one pretend that one had come anywhere near a 'successful' modelling of the 'life' or 'place' of some individual ? But such material is not at all easy to collect, and presumably there is an end to one's patience or persistence ? . . . . . (E,C1)

All that may be modestly claimed of the present arrangement is that it moves things in somewhat the appropriate direction: away from summary representations which beg all sorts of sociolinguistically relevant interpretive questions, and towards a place where, though the banal is all too obvious, it is at least recognisable as human (E,C2).

(Note: Throughout NC refers only to missing data, unless otherwise specified.)

1. Cityness of informant.
[multiple coding: '5-years' criterion for mobile informants]

/1 conurbation	/2 big town	/3 market town	/4 other	/5 NC
Tyneside	e.g.	e.g.		
Teeside	Bristol	Grantham		
Merseyside	Nottingham	Hexham		
Clydeside	Leicester	Taunton		
London	Swansea	Shrewsbury		
Manchester	Edinburgh	Cambridge		
Birmingham	Cardiff			
Sheffield	Chelmsford			
Leeds	Peterborough			
Stoke	Reading			
Solentside	Oxford			
Belfast				
Dublin				

Comment:

Whether a person views it positively or negatively, the extent to which a person is citified (or citybound) is often of far-reaching importance both to that person and to those who perceive that person - whether newly, or familiarly. Such considerations motivate this criterion. In terms, for example, of being in a city, a positive view might emphasise culture, opportunity, variety, education, mobility, anonymity, . . . , as being relevant. A negative view might emphasise alienation, pollution, exploitation, corruption, unrealistic economics, debt traps, . . . , as being relevant. It seems reasonable to see the positive and negative views as connecting rather than separating those who seem to choose to, or seem to be forced to, live in cities.

Of course the apparent finiteness and inflexibility of the lists associated with each state look as bad as an approaching straightjacket. Worse still, it's perfectly possible for someone who

lives in a conurbation like Tyneside to feel that they do **not** live in a conurbation - which must, after all be someone else's word for 'it' - but in a 'village', a 'frazzione', a 'market town', having unique characteristics invented for 'it', invested in 'it', by each one who is living 'there'.
[Let us not forget the informant living in Byker (a five minute eastward bus ride from the centre of Newcastle) who had **never** been to North Shields which is more or less the nearest beach (perhaps some seven miles to the east) and who thought that a journey to Newcastle was a huge expedition which she had only made eight times in her life - each of which she could remember. At the time of talking, she was fifty five. This unwillingness to grant the wider horizon may not be as unusual as we think. It is surely a centrally critical problem for the notion sociologically homogeneous group that a person can be as different from another 'member' of the same group, as a striking miner is from a Tory Minister of Energy, or as an academic is from a Trappist.]

But notice that the lists are only illustrative guides for the states of this criterion, they are not defining specifications. Nevertheless, you may object, either this is a criterion which reflects the informant's subjective sense of cityness, or it is a quasi-objective specification by the researcher, or the researcher's coding frame, of the informant's cityness. But I can see no very good reason why these two cannot be in an amiable, if ill-defined, resonance (Zadeh 1965).

2. Regionality of informant.

[multiple coding: '2-years' criterion for mobile informants]

/1 UK Northern /2 UK E. & W. Ridings /3 UK North West
/4 UK North Midland /5 UK Midland /6 UK Wales /7 UK Eastern
/8 UK London & South East /9 UK South /10 UK South West
/11 UK Scottish Lowlands /12 UK Scottish Highlands
/13 UK Ulster /14 Republic of Ireland /15 N. America
/16 Antipodes /17 South Asia (Indian sub-continent)
/18 Hamitic Africa /19 Germanic Europe /20 Caribbean
/21 SE Asia /22 Arab Africa /23 Romance Europe

/24 Slavic Europe /25 S. America /26 NC

Comment:

A person's sense of their 'world', in spatial terms, can be a source of pleasure, of terror, and of romantic and/or destructive stereotypes, as well as being useful (in getting home, or from A to B). We are not here concerned with the significance (compared to other dimensions which project a particular individual's place in this, or any other stateable, property-space), or the reality, for an individual, of his or her spatio-regional constructions, but rather only with their nameability.

Objection may be made concerning the superficial geographicalness of the states of this dimension. And of course the dissection of the world's surface is undertaken in a very uneven and Britanno-centric manner. But then both of these biases are ones which are likely to be reflected in the spatial feelings and representations of the person being interviewed. It is well known, for instance, that a person's representation of her or his space is locally dilated, and non-locally constricted - varyingly, according to usage, belief and commitment

In spite of these relativities, the extent to which people overlap in their regionality will be, proportionately, a considerable source of mutual interest and inference.

3. Regionality of both parents.

[i.e. the responses for both parents are coded on one dimension; multiple coding: '2-years' criterion for mobile parents; states as in preceding dimension]

/1 /2 /3 /4 /5 /6 /7 /8 /9 /10 /11 /12 /13

/14 /15 /16 /17 /18 /19 /20 /21 /22 /23 /24 /25 /26

4. Number of moves per five year period before marriage.

/1 0 /2 1 /3 2 /4 3

/5 4 /6 5 /7 5+ /8 NC

5. Number of moves per five year period after marriage.

/1 0 /2 1 /3 2 /4 3

/5 4 /6 5 /7 5+ /8 NC

Comment:

The disparity, or non-disparity as the case may be, between dimensions 4. and 5. incorporates a good deal of information about the aspirations and consequences of any marriage. Of course the motives for the moves may be generated within the marriage, where they may be either cooperative or unilateral, or they may be imposed from without (family, landlord, employer, etc.)

6. Age.

/1 17-20
/2 21-30
/3 31-40
/4 41-50
/5 51-60
/6 61-70
/7 71-80
/8 81+
(/9 NC)

Comment:

These arbitrarily equally bounded states cannot hope to encompass all the many currents of feeling that individuals have about their own age and about the ages of others. The details of emotional and authority relations, both positive and negative, cut across the many and varied stereotypes of ageism and youthism, which though they are infuriating are of very wide currency. In addition, it is important to remember that the age group of an individual never constitutes a necessary and sufficient generator of that individual's behaviour.

7. Gender.

/1 M
/2 F
(/3 NC)

Comment:

In modern urbanised populations, whether in developing, nondeveloping or de-developing nations, it is arguable that gender is the single most significant variable underlying the self-constructions, attitudes, values, aspirations, achievements, and behaviours of individuals, and consequently underlying the behaviours amongst individuals. This has generally not been thought to be true by the many at large (but see e.g. Rich 1980), or by many in sociolinguistic studies (but see e.g. Horvath 1985; Lesley Milroy 1989).

8.School leaving age.

/1 before legal minimum /2 legal minimum /3 +1 /4 +2
/5 +3 /6 +4 /7 +5 /8 NC

[Legal minima:	Age	> 92	: chaotic/ non-existent
		77-92	: 13
		52-76	: 14
		< 52	: 15]

9. Tertiary and further education.

/1 none /2 full-time university &/or polytechnic
/3 full-time technical/ nursing/ secretarial
/4 full-time college of education /5 block release
/6 day release /7 night school
/8 self-taught & correspondence /9 NC

10. Attitude to education of self.

[multiple coding]
/1 negative /2 basic skills (RRR) /3 liberal /4 job-oriented
/5 liberal and job-oriented /6 NC

11. Attitude to education for offspring.

[multiple coding; states as above]

/1 /2 /3 /4 /5 /6

12. Distinction between education of boys and girls.

/1 yes /2 no /3 NC

13. Positive distinction between parental and school roles

/1 yes /2 no /3 NC

14. Parental control of children.

[multiple coding]

/1 direct verbal /2 indirect verbal
/3 direct physical /4 indirect physical /5 NC

15. Marital status.

[multiple coding]

/1 married /2 single /3 divorced /4 separated
/5 widowed (/6 NC)

16. Religion.

/1 active /2 inactive /3 anti /4 NC

17. Nuclear family size.

[i.e. including breadwinner(s) and spouse]

/1 1 /2 2 /3 3 /4 4 /5 5 /6 6 /7 6+ /8 NC

[Note: Unmarrieds living at home with two parents are coded /3; unmarrieds living alone are coded /1; marrieds with no offspring are coded /2; etc.]

18. Sex distribution of offspring.

[absolute numerical; NC includes NA]

/1 zero bias /2 F bias /3 M bias /4 NC

19. Average age gap between offspring.

[NC includes NA]

/1 1 year /2 2 years /3 3-4 yrs. /4 5-6 yrs. /5 7-8 yrs.
/6 9+ yrs. /7 NC

20. Distance of spouse's primary regionality.

[NC includes NA]

/1 same local authority /2 < 50 miles /3 > 50 miles /4 NC

21. Micro-environmental preference in terms of sentiment.

/1 neutral /2 dissatisfied /3 satisfied, ambitious
/4 satisfied, stable /5 NC

22. Micro-environmental preference in terms of housing.

[same states as 21.]

/1 /2 /3 /4 /5

23. Interviewer's assessment of decoration, furnishing and domestic equipment.

(a) 'Taste' aspiration:
/1 good /2 bad /3 indifferent /4 NC
(b) Financial commitment to this taste [10 is high]:

/1 /2 /3 /4 /5 /6 /7 /8 /9 /10 /11 NC

24. Rateable value

[a quantitative variable; x_{max} to be determined from samples]

25. Macro-environmental preference.

[finance and/or occupation no object]

(a)
/1 rural /2 smaller town /3 same size /4 NC

(b)
/1 south /2 north /3 nowhere else /4 NC

26. Positive Tyneside consciousness.

/1 yes /2 no /3 NC

27. Social integration with neighbours (as claimed by informant).

[multiple coding]

/1 non-existent & unknown /2 non-existent & known
/3 antagonistic /4 minimal, pleasant /5 cordial
/6 intimate /7 NC

28. Father's occupation [cf. Hall & Jones (1950)].

/1 professional & high administrative /2 managerial and executive
/3 inspectional, supervisory & other non-manual (higher grade)
/4 inspectional, supervisory & other non-manual (lower grade)
/5 skilled manual & routine non-manual /6 semi-skilled manual
/7 unskilled manual /8 NC

29. Informant's present occupation
(or spouse's if informant is not primary breadwinner).
[states as above]
/1 /2 /3 /4 /5 /6 /7 /8 NC

30. Informant's first occupation.
[states as above]
/1 /2 /3 /4 /5 /6 /7 /8 NC

31. Job preference.

(a)
/1 prospects /2 immediate gain /3 NC

(b)
/1 thinking (new elements) /2 learned (no new elements) /3 NC

(c)
/1 supervised /2 self-deciding /3 NC

32. Job satisfaction

[match between 31 & 29]

/1 3 /2 2 /3 1 /4 0 /5 NC

33. Daily exposure to radio & television.

(a)
/1 predominantly radio /2 predominantly television
/3 radio only /4 television only /5 non-owner /6 NC

(b)
/1 intense, non-selective /2 intense, selective
/3 non-intense, non-selective /4 NC

34. Regular drinking habit/housework as <u>hobby</u>.

/1 yes /2 no /3 NC

35. Leisure satisfaction.

/1 satisfied /2 partially satisfied /3 disgruntled /4 NC

36. Hobbies [examples in brackets].

/1 active, expensive, rule-based, club (rackets)
/2 active, expensive, rule-based, non-club
/3 active, expensive, non-rule-based, club (hunting)
/4 active, expensive, non-rule-based, non-club (DIY, veteran car driving)
/5 active, cheap, rule-based, club (amateur football)
/6 active, cheap, rule-based, non-club (rounders)
/7 active, cheap, non-rule-based, club (cross-country)
/8 active, cheap, non-rule-based, non-club (fellwalking, gardening)
/9 sedentary, expensive, rule-based, club (roulette)
/10 sedentary, expensive, rule-based, non-club (poker)
/11 sedentary, expensive, non-rule-based, club (pedigree breeder)
/12 sedentary, expensive, non-rule-based, non-club (punter)

/13 sedentary, cheap, rule-based, club (billiards/snooker)
/14 sedentary, cheap, rule-based, non-club (scrabble)
/15 sedentary, cheap, non-rule-based, club (amateur drama)
/16 sedentary, cheap, non-rule-based, non-club (reading)
/17 active, expensive, collecting, club
/18 active, expensive, collecting, non-club
/19 active, cheap, collecting, club
/20 active, cheap, collecting, non-club (sea shells)
/21 sedentary, expensive, collecting, club (book/picture clubs)
/22 sedentary, expensive, collecting, non-club (stamps, antiques)
/23 sedentary, cheap, collecting, club
/24 sedentary, cheap, collecting, non-club (newsworthy faces, Green Shield stamps (?))
/25 NC

37. Connection between occupation and voting behaviour.

/1 approve /2 accept /3 disapprove /4 NC

38. Voting preference at last election.

/1 Conservative /2 Labour /3 Liberal /4 other /5 Communist
/6 refusal to answer /7 floater /8 NC

We may briefly examine the fate of the dimensions of the Social Space (SSp) as they respond to the same sample of informants whose effect upon the phonological subspaces of the VSp we have already examined (see above, Appx. A, Section 7.) [This summarises Jones (1978 (=1983)), for more details see her pp. 158-193, 247-287).]

There are 52 informants comprising:

(i) a fragment (the two lowest strata) of a stratified random sample from Gateshead (stratum 4 [Rent= £4+ per week]: 11 informants; stratum 5 [Other council house]: 34 informants; total 45 informants), and

(ii) a fragment of a random sample (drawn by the clustering method (Moser & Kalton 1971)) from Newcastle (7 informants). (These subsamples are not statistically compatible, but then, no population inferences are being made.)

It is possible that a classification of the informants upon the dimensions of the SSp will simply reproduce as clusters these three subsamples, but, contrary to Jones (1983: 165), it does not seem reasonable to entertain this as a higher-valued expectation.

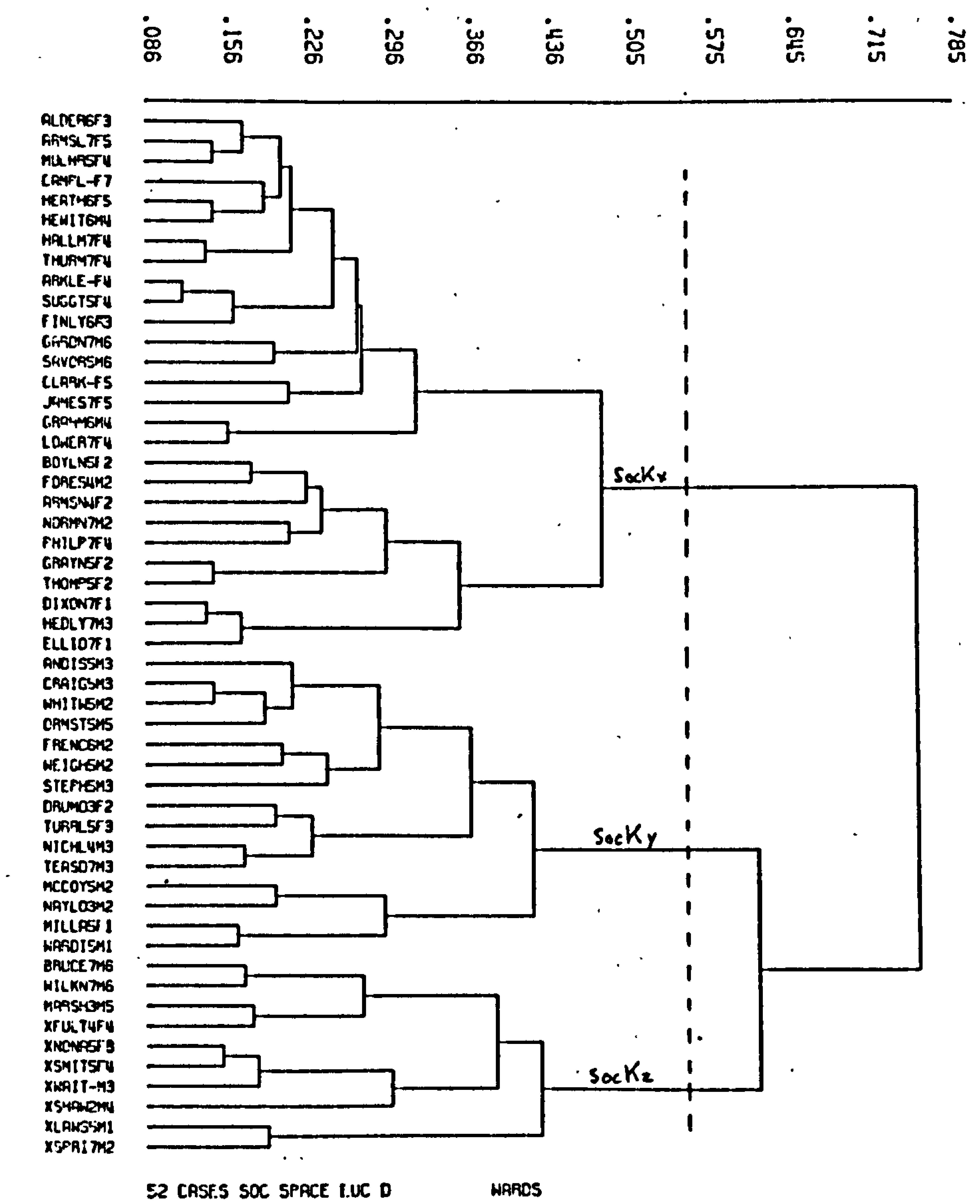
The sample of 52 informants consists of 25 females and 27 males, distributed as to age, education and occupation as follows:

(Abbreviations: lma = legal minimum school leaving age (varies according to age of informant); fe = Further Education; tech = Technical College; ft = fulltime; acad. = academic; manag. = managerial; exec. = executive; inspc. = inspectional; spv. = supervisory; sk.= skilled; mn. = manual. The categories within any heading are mutually exclusive.)

Age		Education		Occupation	
1. 17-20	10%	a. lma, no fe	63%	2. manag./exec.	2%
2. 21-30	25%	b. lma+, no fe	0%	3. inspc./spv. (high)	6%
3. 31-40	19%	c. fe - tech	25%	4. inspc./spv. (low)	8%
4. 41-50	23%	d. ft higher	6%	5. sk. mn	40%
5. 51-60	12%	e. ft acad.	6%	6. semi-sk. mn.	13%
6. 61-70	8%			7. un-sk. mn.	31%
7. 71-80	2%				
8. 81+	2%				

The sample was clustered using the CLUSTAN 1A clustering program package (Wishart 1969), in respect of the dimensions specified above. The CLUSTAN options which were used were as follows: the comparison function was binary Euclidean distance (D), and the clustering algorithm was Ward's (minimisation of the Error Sum of Squares). (See above, Appx. D, Sections 2 &3, for discussion of comparison functions and clustering algorithms.) The dendrogram overleaf gives the sample clustered on the dimensions of the SSp. The number of SSp clusters in the sample is taken to be three (the 3-K level), since plotting the number of clusters obtained against increasing values of D shows the first plateau to be at the 3-K level (or, put another way, there is a relatively large jump in the D value from the point at which 6-K fuses to 3-K to the point at which 3-K fuses to 2-K (see dendrogram)). But there is also a significant plateau at the 2-K level. Nevertheless, because the two clusters which would fuse to make 2-K from 3-K are very different as to their distribution

558C



of age, education and occupation attributes, the 3-K level is preferred to the 2-K. The clusters (see dendrogram) are named SocKx (Social cluster x), SocKy, and SocKz, and they consist of 27, 15 and 10 members respectively. An analysis of the three clusters in respect of age, occupation, sex and education shows the following.
Percentage deviations of age distribution in the three clusters from sample expectations are:

	SocKx	SocKy	SocKz
Age gp.			
17-20	- 2.5	+ 3.0	0.0
21-30	- 3.0	+ 15.0	- 15.0
31-40	- 8.0	+ 21.0	- 9.0
41-50	+ 10.0	- 23.0	+ 7.0
51-60	+ 3.0	- 5.0	- 2.0
61-70	- 0.5	- 8.0	+ 12.0
71-80	+ 2.0	- 2.0	- 2.0
81+	- 2.0	- 2.0	+ 8.0
Educ. cat.			
a	+ 15.0	- 30.0	+ 4.0
b	0.0	0.0	0.0
c	- 6.0	+ 29.0	- 25.0
d	- 2.0	- 6.0	+ 16.0
e	- 6.0	+ 7.0	+ 5.0
Occup. gp.			
2	- 2.0	- 2.0	+ 8.0
3	- 6.0	+ 7.0	+ 4.0
4	+ 0.5	- 1.5	+ 2.0
5	- 15.0	+ 27.0	- 10.0
6	+ 8.0	- 6.5	- 13.0
7	+ 15.0	- 24.5	- 1.0
Sex distribution (male ratios)			
	8/17	12/15	7/10
	(30%)	(80%)	(70%)

It is important to note that in most cases, for each of these social indices, each SC has one modal tendency, but that there is also significant mixing. SocKx tends towards the middle aged groups (41-60), is overwhelmingly female, tends towards low educational status and low occupational status. SocKy tends towards youth

(17-40), is predominantly male, further education tends to be vocational/technical, and skilled manual & routine non-manual occupations are relatively highly represented. SocKz, also predominantly male, shows a mixed age distribution, shows two educational extremes, and a relatively high distribution of supervisory and managerial occupations. For SocKx and SocKy, we may assume, to a certain extent (see figures above), that responses upon other dimensions of the SSp are dependendent upon age and sex. For SocKz no such assumption can be made. The diagnosticity of these other responses for each group are as follows. Diagnostics are listed in decreasing order of diagnostic power. The ratios after each diagnostic indicate its frequency in the K over its frequency in the sample.

SocKx

(Binary % freq. ratio values: positive diagnostics in the range 1.93 to 1.30. [Jones (1978) also gives the negative ones.] In spite of the level of diagnostic power, the underlined diagnostics are those of most interest.)

Leisure satisfaction=disgruntled (1/1); hobbies=7 [see list above] (2/2); hobbies=22 (2/2); hobbies=15 (2/2); citiness=market town (1/1); dist. spouse<50m.>local authority (8/8); citiness=other (2/2); parents' reg.=UK N.Midland (2/2); parents' reg.=UK Midland (1/1); parents' reg.=UK Lowland (1/1); TV/radio=predom. radio (1/1); housework as hobby (8/9); occup./voting behav. connection (8/9); father's occup.=7 (12/14); info's 1st. occup.=7 (17/20); info's pres. occup.=6 (5/6); 'taste asprn.'=indifferent (20/25); vote Lab. (19/24); distinctn. educ. boys/girls (6/8); integr. neighbours non-exist./known (3/4); micro-env. (housing) dissatisfied (9/12); sex bias of childr.=M (9/12); info's pres. occup.=7 (11/15); parental control=dir. phys. (16/22); parental control=indir. phys. (5/7); integr. neighbours non-exist./unknown (5/7); TV view=intense, non-selective (7/10); parental control=dir. vbl (7/10); + Tyne consciousness (17/25).

Overall, this group is lower working class. However, it is far from homogeneous. Responses to the questionnaire are not stereotypical, and non-modal occupations and educations are present.

SocKy

(Binary % freq. ratio values: positive diagnostics in the range 3.47 to 1.50.)

Info's last occup.=3 (1/1); financial commit. taste=10 (2/2); fe=coll. educn. (1/1); lma+3yrs. (2/2); parents' reg.=UK Lowland (1/1); parents' reg.=UK E&W Ridings (1/1); lma+5yrs (1/1); fe=day release (4/5); lma+2yrs. (3/4); hobbies=12 (3/4); att. educ.(childr.)=RRR (3/4); taste asprn.=good (11/15); occup.=3 (2/3); vote=Con. (4/6); integr. neighbours=antag. (3/5); vote=floater (3/5); macro-env. pref.=south (3/5); drinking as hobby (7/12); fe=night sch. (5/9); parental control=indir. vbl (5/9); lma+1 yr. (6/11); financial commit 'taste'=6-7 (7/13); occup.=5 (10/19); info's 1st occup.=5 (10/19); hobbies=4 (1/2); integr. neighbours=intimate (1/2); job pref.=interest (8/16); reg.=UK Lond/SE (1/2); voting pref.=refusal (1/2); fe=univ/poly (1/2); TV/radio=non-owner (1/2); sex bias childr.=Ø (4/8); father's occup.=3 (1/2); father's occup.=4 (2/4); TV only (1/2); father's occup.=5 (8/17); attit.educ=liberal (7/15); sex=M (12/26); no. moves in 5 yrs. after marr.=1 (5/11); TV view=non-intense, non-selective (10/22); disapprove connec. occup./vote (4/9); attit. educ.=job oriented (7/16); dist. spouse's reg.=same loc. auth. (10/23).

Overall, this group has higher job-satisfaction, less class-dominated politics, and more positive attitudes to education than SocKx.

SocKz

This cluster has a much wider range of values for any given social dimension than SocKx and SocKy, or, put another way the intra-cluster distances between its members are much higher. Consequently the diagnostics, though more numerous than for the two denser SCs, are much less significant.

It will be clear that none of these three SCs are anything like monothetic groups (i.e. those relying on a set of necessary & sufficient criteria to establish group membership).

We may now examine the interaction between the fate of the **social** dimensions which we have just discussed for this sample, and the fate of the VSp **phonological** dimensions in the face of the same sample discussed above (Appx. A, Section 7, p.120 ff.).

Superimposition of Figure A15 on Fig. A10 (in pocket at the back of Vol iii) shows the distribution of the sample across the three clusters (SocKx red, SocKy green, SocKz brown), having respectively 27, 15, & 10 members. Superimposition of Fig. A11 on Figs. A15 & A10 shows the social distribution of the VCs of speakers who are similar in respect of monophthongs (K1, K2, K3). K3 is exclusively composed of members of SocKz, but note that there are 3 members of SocKz who are not in K3. Similarly, superimposition of Fig. A12 on Figs. A15 & A10 shows the social distribution of the three VCs representing similarities between speakers in terms of diphthongs and reduced vowels. And Fig. A13 on Figs. A15 & A10 does the same for consonants.

The derived VCs of Fig. A14 are those in which membership across all three phonological subspaces is maintained; that is, they group informants who are in K1 and KA and Kα, and then those who are in K1 and KB and Kα, and so on. Superimposition of Fig. A14 on Figs. A15 & A10 shows that the relationship between the eight derived VCs and the SCs is no more straightforward than the relationship between the SCs and any one of the VSp subspaces (represented by Fig. A11, or Fig. A12, or Fig. A13).

For both the linguistic classification (see above Appx. A, Section 7) and the social classification there were secondary plateaux of increasing values of D against decreasing numbers of Ks. In both cases these were at K=2. But the 2-K level is only an apparent simplification, as can be seen from the following considerations.

At the 2-K level the linguistic classification gives LK1 (Gateshead) and LK2 (Newcastle). LK1 = (KA+KB)+(K1+K2)+(Kα+Kβ). LK2 = (KC)+(K3)+(Kγ). (This can be seen by superimposing Fig. A16 upon Figs. A11, A12, & A13.) At the 2-K level in the social classification the Newcastle informants are **not** separated from the Gateshead informants, but SocKz fuses with SocKy (see the dendrogram above on p. 558C). As Jones (1978: 286-7) writes: "The 2-K level only apparently simplifies the picture of overlap of cluster membership: this is because the number of clusters is smaller, therefore the combinations are fewer. If **n** is the number of clusters (2K or 3K) and **x** is the number of classifications (1 soc. + 3 ling. = 4) then $\mathbf{n}^{\mathbf{x}}$ gives the maximum number of 'derived' clusters of the type (SocKx, K1Aα), (SocKx, K1Aβ), etc. ([This is] a measure of degree of overlap between classifications.) For 3 clusters per classification (3K level), $\mathbf{n}^{x}$ = 71. For 2 clusters per classification (2K level), $\mathbf{n}^{x}$ = 16. [Superimposition of Figs. A10, A14 & A15 shows] that the four classifications superimposed at the 3K level give 15 combinations (derived clusters) out of a possible 71. [Fig. A10 & Figure A16] show that the four classifications superimposed at the 2K level generate 3 combinations, out of a possible 16. The ratios 3/16 (2K) and 15/71 (3K) are quite similar in magnitude (0.1875 and 0.2113 respectively). Thus in terms of the potential number of overlaps between classifications, the situation at the 2K level is approximately as complex as that at the 3K level."

Correspondences in K-membership (3-K level) between social & linguistic spaces [after Jones (1978: 250, Tables 58-61)]

		SocKx	SocKy	SocKz	Tot
	K1	22	10	2	34
%FON1:	K2	5	5	1	11
SocSp	K3	-	-	7	7
	Tot	27	15	10	52
	KA	11	9	1	21
%FON2:	KB	16	6	2	24
SocSp	KC	-	-	7	7
	Tot	27	15	10	52
	Kα	11	11	1	23
%FON3:	Kβ	16	4	2	22
SocSp	Kγ	-	-	7	7
	Tot	27	15	10	52
	K1Aα	4	4	-	8
	K1Aβ	5	4	1	10
Derived	K1Bα	4	2	-	6
clusters:	K1Bβ	9	-	1	10
SocSp	K2Aα	-	1	-	1
	K2Aβ	2	-	-	2
	K2Bα	3	4	1	8
	K3Cγ	-	-	7	7
	Tot	27	15	10	52

Appendix F

On definitiveness, refutation and appropriation

But, when all is, finally, considered - were that conjunction of three projects at all possible - one should surely recall George Eliot's Gyre which not only dissolves Occam's Razor but ensures that any determinate empiricism is self-deluding. In addition, it anticipates quantum physics, the uncertainty principle, chaos theory and deconstruction, and suggests that not only 'progress', but the never-ending accumulation of 'results' which are required to underpin that unlikelihood, are very probably the hopeful fictions of the self-important (F,C1).

Meanwhile, goes her Gyre - regardless of the developmental point-in-time of some analysis, or theory, or re-presentation - **the indefiniteness remains, and the limits of variation are really much wider than anybody would imagine.**

* * *

And - threading Vico's circle on Eliot's gyre - this, from another hand (Lumsden 1989). [N]

[N] My deepest thanks to Bob Lumsden for allowing me to quote in full, as follows, his recent piece on the Bare Bones of Deconstruction.

* * *

Deconstruction, the Bare Bones

In the following, certain key terms, among them Derrida, de Man, and deconstruction, exist as markers in a process of reasoning - thus: "Derrida", "de Man", "deconstruction". No co-relation is to be supposed between these terms and 'actual' theories which these names 'actually' locate.

I accept that Derrida's theory of deconstruction puts in question the stability of all signification, therefore of all communication.

According to Derrida's deconstruction, to understand is to revise.

Thus, Derrida's deconstruction inherently invites revisions.

I treat Paul de Man as the principal revisionist of Derrida .

I treat most if not all of what passes today as deconstructionist as based on Paul de Man's, rather than Derrida's, idea of deconstruction.

I suggest that Derrida's account of language requires that one deal with the question of solipsism, before those explanations commonly used by linguists and literary critics need be considered at all.

Derrida's argument

Derrida's argument begins from de Saussure's premise about the way languages signify. Languages signify in being situated within a series of differences which may be arranged in a hierarchy of increasing complexity - phonetic, phonemic, and lexical (Saussure). There is necessarily a difference between a thing in the world, or a concept, and those sounds used to signify that thing or concept.

The underlying principle of 'differ' and 'defer' which Derrida' gathers into one term, différance, is that of an intractable gap between sign and referent.

'Defer'-ence signifies this gap as it appears between what is understood by a term **at its time of utterance**, and what is **later** understood by it.

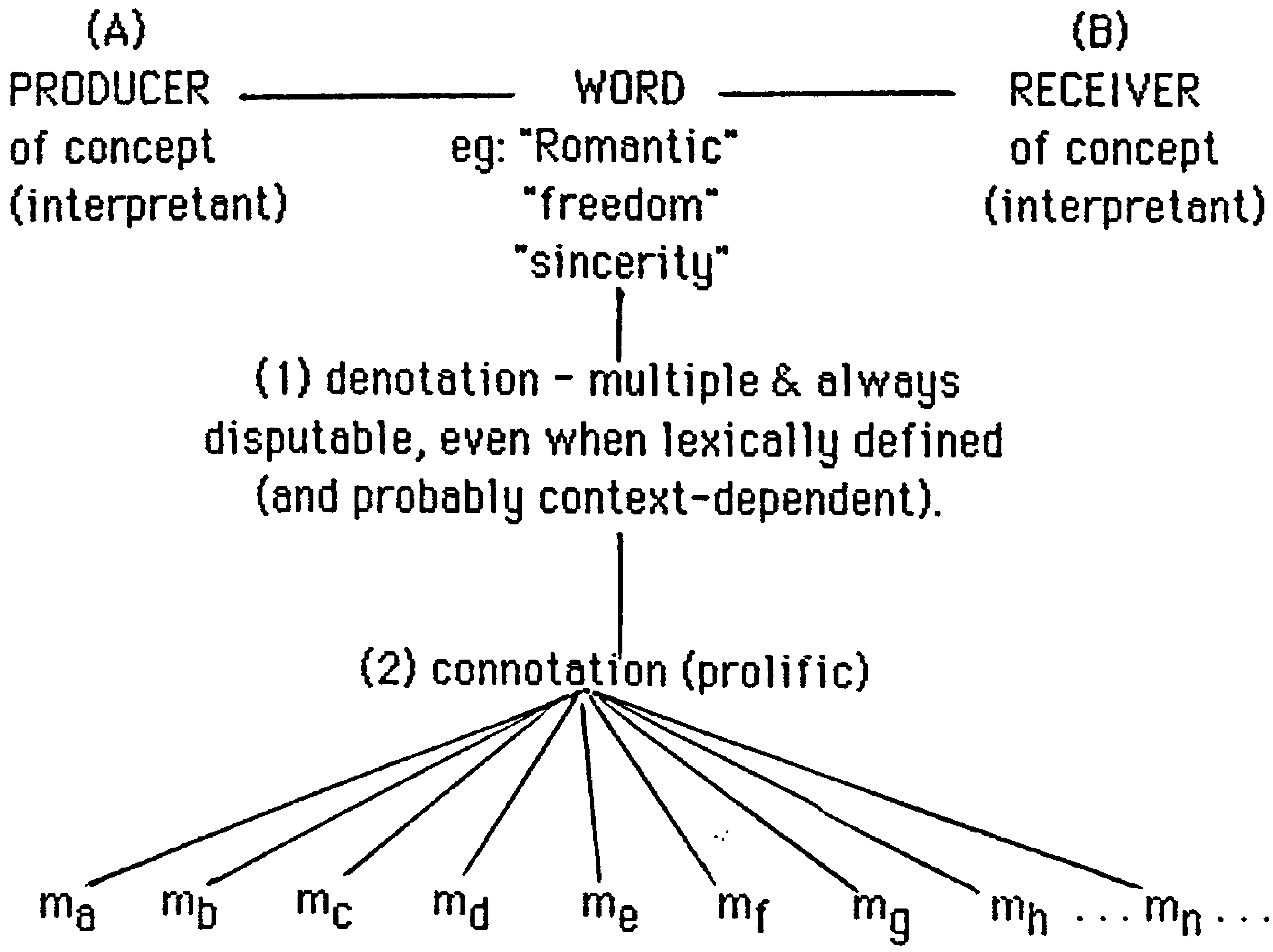
Derrida's (metaphoric) way of stating this, is that every sound-concept is a **trace** of one which has gone before it, and that such a trace itself invites substantiation by another, which also will be a trace of what it represents. And so on, for as long as explanation continues.

'Differ'-ence describes the distance **spatially**, on any signifying matrix, between a sound or word uttered, and all those sounds or words not **present**, but implied by the utterance, and necessary for an understanding of it.

Thus, what is produced in speech or writing depends for its meaning on what is implied, but not realized. Both temporal and spatial discontinuity (defer-ence, and differ-ence) are therefore essential to the functioning of language. Such differences - which I am calling "gaps" - are therefore neither problems to be solved, nor aberrations for which one must wait for a solution. They exist, rather, at the heart of signification.

One way of translating this concept of the essential gap into traditional terms is to consider a favourite literary-critical item, **connotation.**

Connotation, seen from the Derridaean perspective, is a recognition that difference is inevitable, and even fruitful, in the construction of meaning. Consider the following simple diagram:



It is highly unlikely that concept-producer (A), and concept-receiver (B), will be foregrounding or focussing upon the same connotative node of any single word, at any time, let alone the same nexus of that ever-proliferating complex established by subsequent supposedly 'explanatory' or 'qualifying' passages of speaking or writing.

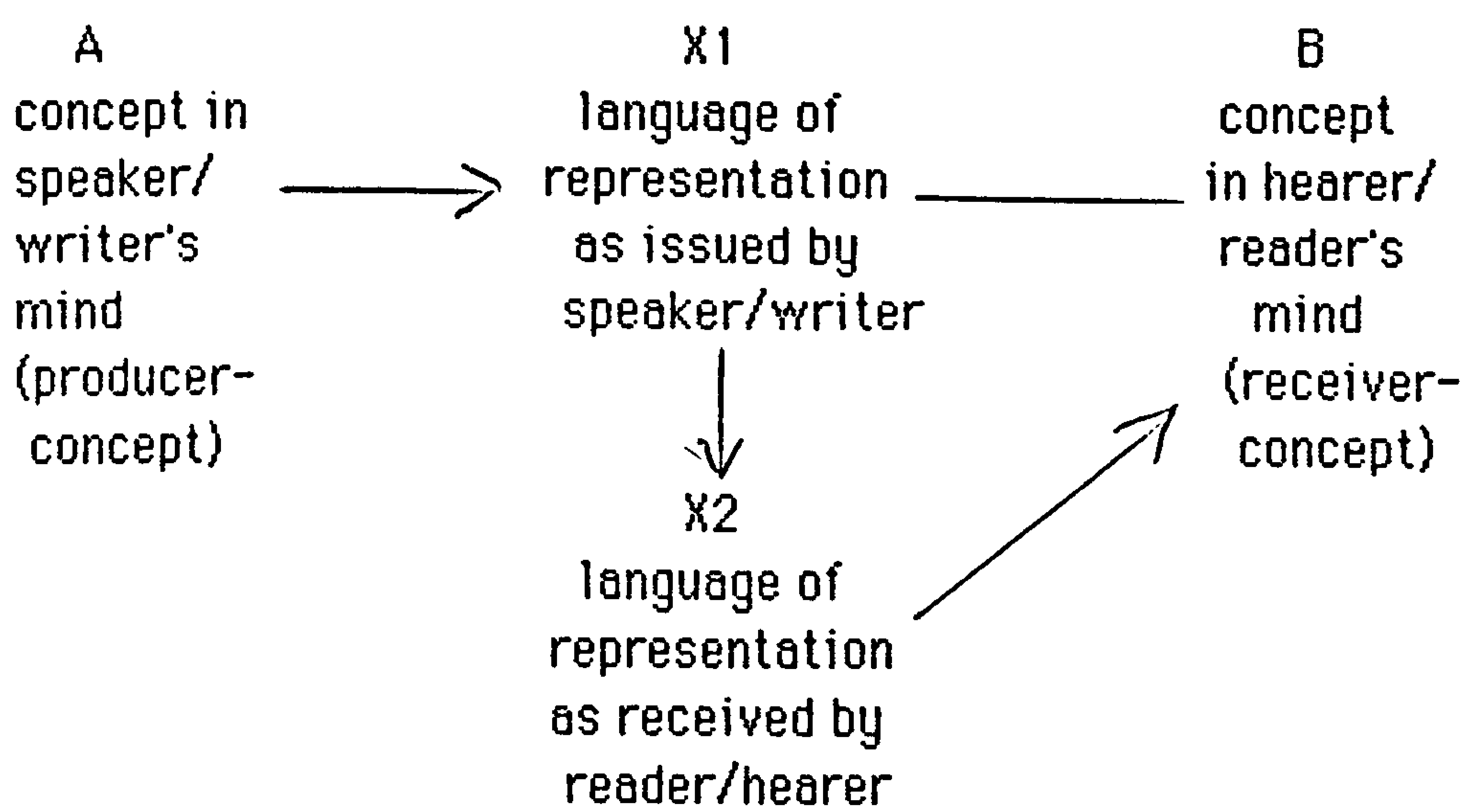
This is to say of 'gaps', in conventional terms, what Derridaean theory says about them in its way, which is:

(a) **Connotation**
it is very unlikely that the concept in the language producer's mind, and the concept in the receiver's mind (whether of hearer or reader) will closely co-relate or,
(b) **Derrida**
it is impossible that the two can correlate.

The isomorphic solution

It is sometimes said that producerconcept and receiverconcept stand in isomorphic relation to that language which passes between producer and receiver. Also, that a producer's (writer/speaker's) concept, and a receiver's (reader/listener's) concept, drawn from the script which they exchange, will also stand in isomorphic relation to each other.

But to suppose this is to reverse the whole trend of Saussurean-Derridaean difference, as described above, and for no apparently good practical reason. Thus, in the following diagram



the isomorphicist supposes that concept A, in a speaker's mind, and concept B, in a listener's mind, mean 'the same' because each of them stands in a similar ratio to its referent.
But this is to assume that the languages of representation, X1 and X2, are congruent, rather than *necessarily divergent.* It is to assume that the Saussurean-Derridaean model (1), described above, is wrong without arguing against it.
Which is, to re-iterate:
since expressions only mean insofar as they are different from those items (concepts, feelings, or objects of reference) which they indicate, all linguistic presentations must be displacements of that original (2).

According to this model, language is the re-presentation of meaning *which we take as* a presentation. Language is the re-presentation of a meaning which we *mis*take as a presentation (3). *Isomorphism would then be merely one variation of such a mistaking.* (It might be described as a version of the ancient theological error of mistaking analogical comparisons for accurate accounts by overlooking those many points of every analogy which do not correspond to their referent.)

Furthermore, language about language is a re-presentation of a re-presentation (4). And we also habitually tend to take these second order re-presentations as 'bodyings-forth' of an original meaning.

De Man's Deconstruction

Derrida presents language as continuously disseminating, after the model described above. There is thus no possibility, in his model, logically speaking, that any linguistic string can signify by retrieving that orginal experience it sets out to describe. The same holds, a fortiori, of any explanation of an explanation. Thus, the one thing we can certainly know of the following statement:

> The feelings of the mind, expressing things naturally, constitute a sort of universal language which can then efface itself (Derrida 1976: 11);

is, that whatever we understand by it, cannot be what it meant for its initiator, Derrida.
Indeed, it cannot even mean what it is supposed to have meant 'originally' (whatever that is conceived to be), *should Derrida attempt to explain it to himself in the next moment after uttering it.*
The same holds true for:

> But although this long struggle in Arnold makes him a more moving figure, the final victory of the Stoic side of him makes him also a little bleak (Lucas 1961: 58);

or,

> 'Justice' was done, and the President of the Immortals, in Aeschylean phrase, had ended his sport with Tess. And the d'Urberville knights and dames slept on in their tombs unknowing (Thomas Hardy, Tess of the d'Urbervilles);

and even,

> Considerable confusion arises when this distinction is not made, when what is really a general definition is in effect presented as though it were a language-particular definition - and this happens quite frequently in traditional grammar, especially traditional school grammar (Huddleston 1982: 2).

We have here, in turn, a description by the arch-undescriber of descriptions; a literary critic's description of a Victorian poet; a novelist's idea of the order of the world; and a linguist's opinion of definitions in grammars. All are irredeemably other than those "things", those concept-strings, to which they refer, as these concept-strings are other than that world they were created to capture. This is the implication of Derrida's reading of language, though not of some accounts of Derrida.

De Man revises all this in his reading practice.

What de Man does is to look for those places in a text where the sense, or argument level, is countermanded by the rhetorical devices or figures of speech used.
These he regards as places where the partiality or blindness of the writer offers the opportunity for insight, because the text tropeically [sic] declares its limitless incompleteness, the endless endurance of the gap between its pretensions to realistic representation and its actual performance.
Once alerted to this essential doubleness, each re-reading becomes an opportunity to join knowingly with that dissemination (displacement) of meaning which is the only true constant in texts of any kind.

But it is clear that de Man has stabilised one level of meaning - the level of sense, or argument - in order that a second level - the figural or rhetorical - can be played off against it.

He is thus arresting dissemination. He is thus retrieving signification for logic and common sense in a way which is not allowed for in Derrida, since Derrida insists that every reading, even of him, will be a new structuring, (therefore a new version of deconstruction) - a re-reading in the most absolute sense.

De Man's deconstruction thus stands as the first revisionist account of the Derridaean idea, and it is this revisionism, not Derrida, which has been adopted by conservationist critics, insofar as they use deconstruction, and by deconstructionist critics, and by that succession of graduate students which has gathered about them.

Typically, a text – Billy Budd, Great Expectations, a Shelley poem – is taken, and what it seems to be saying is matched against its method of presentation. Then the gaps between the two are marked, and one level is used as a commentary upon the other. The figural, for example, is treated as though it were an unconscious revelation of deeper motives, textually inscribed. As though it were a deeper truth given "in other words", literally.
It is clear that this method is thoroughly de Manian; that it arrests dissemination at the level of argument, (by way ofthe figural level); and that, therefore, it is non-deconstructionist, in the Derridaean sense.

So far from de Man "crossing over and exchanging implications in truly Derridaean style" (Norris 1982:100), he rewrites Derrida.

But in this he is very useful to us, because he provides an Ur-example of that imperative to recoup reality through our presentations, and to believe that communication is taking place, which marks all linguistic enterprises, even those most aware of the disseminatory claims, and especially those with a professional interest in establishing language at the centre of human meaning.

A return to first principles

(1) Language signifies by differing from that to which it refers.
(2) Language is a re-presentation of meanings which we take as a presentation.
(3) Language is a re-presentation of meanings which we mistake as a presentation.

(4) One of the principal ways of mistaking the language of re-presentation for the language of presentation in recent years, has been to divide language into strata or levels or streams, set one level as commentary upon another, and suppose oneself more presentational (that is, more 'realistic') for having performed this operation.

This is the method of many a structuralist, both literary and linguistic, of stylisticians, whether sociologically or linguistically oriented, and of many post-1950's literary critics. This strategy of realist retrieval I shall call <u>the de Man prospectus.</u>

(5) But unless propositions (1), (2) and (3) in <u>Derrrida's argument</u> are refuted, the <u>de Man prospectus</u> is incoherent.

And, since the language and literary critical practices of linguists and others, and the view taken of the nature of language itself, are characterized by such logical elements as coherence, consecutiveness, and power of explanation, such an illogicality at the heart of their practices would seem to be very damaging to such practices.

(6) Indeed, unless (1), (2), and (3) above, and all that follows from them, can be refuted, the critically aware, academic language user, if s/he will be consistent, has the following options, and I think, only these:

(a) to try to discover whether communication is actually taking place or whether people just think it is;
(b) to decide that communication does take place, and is language-centred, but that this cannot be demonstrated from within the logic of language;
(c) to decide that communication does take place, but that it is not language-centred.

It may be that supporters of (b) or (c) will find themselves able to draw on a good deal of empirical and logical evidence, both from within and beyond language.

For example, the existence of paradox and contradiction; or the apparent falling away of the need for language among close friends, and people who live together for many years, together with the feeling that essential communication is being enhanced, rather than impoverished, in such circumstances, - these, among other empirical examples might be taken as indications that language is an auxiliary to communication, and not at its centre.

But whatever one may decide, the burden of establishing their practices in that logic to which they aspire, rests, I submit, with those who behave as though there were no important distinction between presentation and re-presentation, and as though their practices were not fatally hurt by their refusal to repair this elementary, foundational, rent in their argument.

Given that their practices are generally so thoroughly grounded in ideals of coherence and consistency, it would seem,
in the absence of a counter-argument,
that the only logical alternative for them would be
silence.

* * *

So things are far worse than I, in various places above, have feared.

Not only is any empirical attempt <u>mis</u>representative of what it purports to show because of the indeterminacy of its sampling methods, and because of the absence of population boundaries, and because of the transfinite nature of any property space (Eliot's Gyre, Herbst (1973), Johnson (1970)) (F,C2).

Not only is it likely that, apart from being wearyingly lengthy, any attempt to model right-hander right-hemisphere processes by left-hemisphere analysis will be thoroughly misleading (F,C3).

Not only must the 'real' nature of hearer inference mostly arise from the **complete** cluster of the hearer's needs and interests which arise and weave and dissolve moment by moment (F,C4).

Not only is it possible, of a given set of objects, to generate an **infinite** number of classifications amongst which there can be no externally valid criterion for choosing 'the best' (F,C5).

Not only is there the risk of the code books being lost,

> The messages and guidelines for order exist only, as it were in sand or are written on the surface of waters. Almost any

> disturbance, even mere Brownian movement, will destroy them. Information can be forgotten or blurred. The code books can be lost. The messages cease to be messages when nobody can read them . . . To be meaningful . . . every regularity must meet with complementary regularities, perhaps skills, and these . . . too are written on sand or the surface of waters (Bateson 1979b: 56).

Not only is it likely that there is far more to the world of spirit than any 'meaning', in the sense of effability, can adhere to (F,C6).

But, the very possibility of refutation completely disappears according to Derrida's position (which is itself therefore irrefutable), since the refuter's case can never make contact with that which it seeks to refute (F,C7).

Nothing more can be said.