# THE NATURE OF PHONOLOGICAL VARIABLES IN SCOUSE

G. Knowles

(University of Lancaster)

## 1 *Introduction*

Scouse is the urban dialect which developed in Liverpool in the middle of the last century as the city was growing by wholesale immigration, especially from Ireland. Scouse might be said to be a Lancashire dialect spoken with an Irish accent. In the present century, it has taken part in general developments in British English, and has been extensively standardized in phonology, grammar and vocabulary.

This paper is based on a study of variation in Scouse (Knowles 1974a), which investigated the relationship of Scouse to Northern English and to British Received Pronunciation -- what it is that makes a Scouser sound like one -- in addition to variation within the dialect itself. The emphasis here is much less sociological than in other urban dialect surveys, and any rudimentary social analysis of the variables was undertaken not for its own sake, but to explain the variation.

Scouse variables are rather different in kind from those described by socio-linguists. Labov (1966:49) selects variables which can be quantified on a linear scale; if there are any like this in Scouse, they are trivial, such as /ɑ/ in *park* which varies from front to back of centre in the open position. More commonly, several scales are involved in any variable. There is no single linear scale for the vowels [ɵ: ë: ɛ̈: ɜ: ɛ: ɛə] which can occur in *square*. Scouse /t/ varies in several ways, including (i) degree of velarization, (ii) tip or blade as active articulator, (iii) duration and force of aspiration, (iv) complete or incomplete closure, (v) sudden or delayed release. Such variation cannot be accommodated on a linear scale without gross distortion of the data.

In fact, the familiar 'variables' of the literature are complexes of component variables. For example, Labov's (eh) (1966:51-3) and Trudgill's (a) (1974:85-6) vary in duration, diphthongization and vowel height. The variants are plotted on the scale of vowel height, because this conveniently puts them in the required social order. In other words, the point to be demonstrated is anticipated in the very way the problem is presented.

If the component variables are separated out, the scope of the analysis is changed. We no longer get half a dozen or so quantifiable shibboleths which occur intermittently in speech, but a large number of variables which are difficult or impossible to quantify, and which occur continuously. Now if speech were composed phonetically of strings of segments, it would be reasonable to expect whole segments to be the units of variation. But since segments are realized as carefully co-ordinated and overlapping movements of the vocal organs, it is more realistic to trace variables to their source in speech production.

A problem raised by variation is how to incorporate it in linguistic theory, and in particular into a generative grammar. Labov later presents his variable (eh) as a variable rule (1972:134ff). The difficulty with a generative rule is how to interpret it: it could (a) represent the speaker's knowledge in some way, or (b) model his actual speech behaviour, or (c) symbolize a sound-change, or (d) merely serve as a convenient device to handle the data.

Variation among speakers possibly reflects differences in competence: different age-groups and social classes may have learned slightly different languages. Inconsistencies in a single speaker are a matter of performance: even if he switches codes his knowledge of the language to which those codes belong remains the same. Performance is of course much more than a rag-bag of errors and miscellaneous inexplicables. There must be very precise rules which the speaker uses in performance, and variation can give a good idea of what some of these rules must be like.

In this paper I shall discuss a number of Scouse shibboleths, and demonstrate that variation is not as simple as socio-linguists have made out. I shall then suggest some of the many ways variables can arise in speech production. But first let us consider some of Labov's New York variables which have their counterparts in Liverpool.

1.1 (*r*) Labov treats his variable post-vocalic /r/ (1966:50) as simply the presence or absence of constriction. This sounds fine on paper and in the classroom, but it is totally inadequate to deal with rhotic dialects, at least in this country. Take Belfast speech for instance. For some people, the /r/ will retroflex the preceding vowel and a following alveolar consonant: they certainly have 'constriction'. Ulstermen who speak RP will have no /r/ and no 'constriction'. A number of well-to-do middle-class people will produce the minutest auditory effect corresponding to /r/: it is quite arbitrary at what point this effect is said to lack constriction and therefore not to be an /r/ at all. Or take North Lancashire speech, which at first seems to have lost post-vocalic /r/. It is not uncommon to hear pronunciations which have no constrictions corresponding to /r/, but in which the quality of neighbouring segments can only be explained if /r/ is assumed to be there. Such 'indeterminate cases' Labov would presumably leave out of his calculations. But it is a strange science which dismisses crucial data which happens not to fit the theory!

Scouse has lost the post-vocalic /r/, but a relic survives in the vowel [ɐ:] as in [lɐ:n, ʃɐ:t] 'learn, shirt'. This vowel has a slight raising of the tongue tip as for /r/ which distinguishes it from [ɜ: ë: ɛ̈:] produced with a flat tongue surface. However, tip raising for [ɐ:] gives the vowel slight auditory rounding rather than r-colouring. Depending on whether the analysis is based on articulation or auditory quality, [ɐ:] can be said to realize /r/ or not.

Before /r/ can be used as a variable, its realizations have to be studied in detail. The values of its variants cannot be decided on as a matter of course in advance.

1.2 (*th*) Labov's (th) has the three variants fricative, affricate and stop (1966:55). Scouse has the standard fricative varying with sounds deriving from an Anglo-Irish dental stop which we can symbolize as [T].

Phonetically, any stop has a period of affrication; if it is short enough, it will indicate the place of articulation of the stop without being perceived as a separate segment. The point at which it is useful to distinguish 'stops' and 'affricates' depends on how a particular dialect affricates stops anyway. Scouse [ T ] in *thin* might be heard as an affricate just like any other Scouse initial stop. Scouse post-vocalic stops may have incomplete closure and thus sound like fricatives. Accordingly, [ T ] in *bath* may be mistaken for [ θ ]. At the same time it sounds like a stop on account of its influence on the duration of the preceding vowel. The combined impressions of stop and fricative could even suggest it is an affricate. Clearly, any quantitative analysis is of doubtful value unless a thorough phonetic analysis has been carried out first.

## 2 *Shibboleths*

Shibboleths are the gross acoustic variables which most readily suggest themselves for socio-linguistic analysis. They can be used very effectively to demonstrate that speech varies according to social group and situation; but for any theory which tries to explain variation, or say what it is that is varying, shibboleths have to be broken up into their components.

2.1 (ʊ) In Liverpool, as in other Northern towns, *mud* rhymes with *good* in traditional popular speech, but varies to [ mʌd ] in educated middle class speech under the influence of RP. There are a number of intermediate lower middle class forms that can be transcribed collectively with schwa. The scale for (ʊ) then looks perfectly simple:

(ʊ - 1) : [ ʊ ] (ʊ - 2) : [ ə ] (ʊ - 3) : [ ʌ ]

We can safely predict higher scores for this variable in middle class (MC) than in working class (WC) speech, and in formal rather than informal style.

In fact, [ ə ʌ ] do not pattern like other vowels. For most vowels it is easy to find some kind of norm, especially if they exist as words in isolation, e.g. /u/ 'who', /ɪ/ 'it', /ɔ/ 'or'. Intrinsic allophones, varying in duration and quality, can be explained by the phonetic or phonological environment. Forms of [ ə ʌ ] vary in height, retraction and rounding, but not according to any apparent rules, and without any apparent norm.

The problem lies in our presentation of the variable, as a set of alternative states or postures. What is happening, surely, is that speakers modify their [ ʊ ] without being altogether clear what to modify it to. Whether this modification is detected depends on the ear and socio-linguistic background of the hearer. If sufficient modification is made, then at some arbitrary point the vowel will be heard as different from [ ʊ ]; lesser modification is likely to go unnoticed. Suppose a Liverpudlian makes the vowel of *mud* slightly more open and less rounded than the [ ʊ ] of *good*. This minute difference may be enough for another Liverpudlian to realize that he is 'speaking nicely'; a Southerner is unlikely to notice any difference at all.

2.2 (ɑ) A word like *aunt* has the vowel [ a ] in Scouse to rhyme with *ant*; RP has [ ɑ: ] to rhyme with *aren't*. Some MC speakers adopt the RP vowel, but more commonly a compromise is found intermediate in quality or duration between [ a ] and [ ɑ: ]. We can represent the compromise vowel ad hoc by means of the symbol [ A ]. The variants of (ɑ) seem to lie on a linear scale:

(ɑ - 1) : [ a ] (ɑ - 2) : [ A ] (ɑ - 3) : [ ɑ: ]

Higher scores can be safely predicted for the MC and more formal styles.

However, this variable cannot be taken in isolation, because [ a ] does not occur in RP. Where [ a ] does not correspond to RP [ ɑ: ] it corresponds to RP [ æ ]. In other words, Scouse [ a ] is always subject to modification. There are two rules involved:

(1a) a → æ as in [ bæd, dʒæm ] 'bad, jam'
(1b) a → ɑ: as in [ fAst, fɑ:st; dAns, dɑ:ns ] 'fast, dance'

Some MC speakers -- and I am one of them -- who modify [ ɑ ] and who sometimes carry out rule (1a), will be very reluctant to carry out rule (1b). What do they do with words like *fast, dance*? In most cases [ a ] will be left unmodified, or a slight compromise is found. Very occasionally, rule (1a) is applied, to make the hypercorrect [ fæst, dæns ]. Since this [ æ ] is used, it must be added to the scale for the variable, and phonetically it must be ordered before (ɑ - 1). Thus (ɑ - 0) : [ æ ].

Unfortunately the linear scale puts the variants in the wrong order for computing variable scores. A MC person using (ɑ - 2) and (ɑ - 0) could end up with an average similar to a WC person using only (ɑ - 1).

2.3 (ɜ) In Liverpool and neighbouring parts of the North West, no distinction is made in popular speech between /ɛə/ and /ɜ/, so that *fair* sounds exactly like *fur*. Some MC speakers adopt something like RP [ ɛə ] where appropriate, but more common is a compromise vowel [ ɛ: ] without the centring off-glide. We can set up the variable (ɛə):

(ɛə - 1) : [ ɜ: ] (ɛə - 2) : [ ɛ: ] (ɛə - 3) : [ ɛə ]

Again, we have not so much the choice of three states, but the application of a modification rule. The difficulty arises that the base form (ɛə - 1) is not a single phonetic form, but is itself a variable. We can distinguish three main types: (a) [ ɐ: ], (b) [ ɛ̈: ], and (c) [ ɜ: ]. The type [ ɛ̈: ] has itself a closer variant [ ë: ]. Thus:

(ɛə - 1) : (a) [ ɐ: ] (bi) [ ɛ̈: ] (c) [ ɜ: ]
(bii) [ ë: ]
(ɛə - 2) : [ ɛ: ]
(ɛə - 3) : [ ɛə ]

The type [ ɐ: ] is widespread in the North West, and competes with the Scouse vowels [ ɛ̈, ë: ]. In the WC Roman Catholic community in Liverpool which was surveyed, the Scouse vowels were used almost exclusively; two retired Protestants who lived just beyond the community boundary and happened to get into the sample, both used [ ɐ: ]. This may or may not be significant. Of

the Scouse vowels, the closer variety [ ë: ] is the more advanced, and used by a younger age group. The RP-type vowel [ ɜ: ] is used by the MC, and shades off with slight rounding into [ ɐ: ] and with fronting into [ ë: ].

2.4 (uə) *and* (oə) As in urban speech in England generally, the phonemic merger of /oə/ and /ɔ/ is being closely followed by the merger of /uə/ and /ɔ/. To get just a rough idea of the time scale, we can compare the words *sure*, *shore* and *Shaw* for three age-groups:

| | *born before 1918* | *born 1919-1938* | *born since 1938* |
|---|---|---|---|
| *sure* | ʃuə | ʃuə | ʃɔ |
| *shore* | ʃoə | ʃɔ | ʃɔ |
| *Shaw* | ʃɔ | ʃɔ | ʃɔ |

We can treat (uə) and (oə) together or separately:

(uə - 1) : [uə] (uə - 2) : [oə] (uə - 3) : [ɔ:] *or* (uə - 1) : [uə], (oə - 1) : [oə] (uə - 2) : [ɔ:], (oə - 2) : [ɔ:]

In either case the interesting fact is that the centring diphthongs are not single phonetic forms, but are themselves complex variables realizable by whole sets of phonetic forms. Members of the sets are generated by variation in low level rules.

Take /uə/ for example. A MC person saying /ʃuə/ may carry out the RP rule laxing /u/ before another vowel, and thus produce [ ʃʊə ]. A WC Scouser is more likely to front /u/ and say [ ʃʉə ]. Scouse also has diphthongal forms of /u/ which can be stressed on either element: these produce [ ʃʉ́ə ] and [ ʃʊ́ə ]. In a sequence of three vowels in Scouse, the second readily becomes a glide if unstressed: thus [ ʃʊ́ə ] becomes [ ʃʊ́wə ]. Final schwa in Scouse can be fronted in the direction of Cardinal 3: this produces a host of forms -- [ ʃʉɛ, ʃʊ́ɛ, ʃʉ́ɛ, ʃʊ́wɛ ] and even the hybrid [ ʃoɛ ]. In this way there are at least ten variants of (uə - 1).

In the same way, /oə/ can occur in several phonetic forms, including [ oə, oɛ, ɔwə, ɔwɛ ] in addition to the RP type [ ɔə ].

2.5 (*ng*) In Liverpool, and in other parts of the North West as far south as Birmingham and as far east as Sheffield, [ ŋ ] is for many speakers an allophone of /n/ occurring before /g, k/. Phonological /g/ is pronounced [ g ] at the end of a word, e.g. [ sɒŋg, θɪŋg ] 'song, thing', and also between vowels, as in [ sɪŋgə, baŋgən ] 'singer, banging'. In consonant clusters, [ g ] tends to be deleted, cf. [ rɪŋz, baŋd ] 'rings, banged'. This looks like the simplest of all possible variables, with presence or absence of [ g ]:

(ng - 1) : [ŋg] (ng - 2) : [ŋ]

As it happens, both variants are themselves variables, realized by a set of phonetic forms. This case is more complicated even than the (uə) variable, for the realization sets actually overlap. Many examples cannot be identified without reference to low level rules.

To begin with, the variable has been wrongly presented. It is not

a phonetic variable -- presence or absence of the final velar stop -- but a phonological one, the presence or absence of phonological /g/:

(ng - 1) : /ng/ (ng - 2) : /ŋ/

In (ng - 1), the final voiced consonant has a lengthening effect on the preceding segment, in this case the (velar) nasal, producing [ŋ:g]. Now a lengthened velar nasal uniquely indicates a following voiced velar stop, so the final segment is actually redundant and can be missed off, leaving [ŋ:]. In (ng - 2), the final voiced consonant has a lengthening effect on the preceding segment, in this case the vowel, producing [ɪ:ŋ] in a word like *sing*. Now a very common feature of final nasals in Scouse is that the release of the articulation is actually audible, and sounds like a weak stop: the release of [ŋ] sounds exactly like a weak final [g]. Thus we have in, for example, *sing*:

(ng - 1) : /ng/ : (a) [sɪŋ:g] (ng - 2) : /ŋ/ : (a) [sɪ:ŋᵍ]
(b) [sɪŋ:] (b) [sɪ:ŋ]

This is a rather nice example, because the data are divided up differently according to whether 'presence or absence of *g*' is interpreted phonetically or phonologically. The distinction is significant, because 2(a) [ŋᵍ] is Scouse and non-prestigious, whereas 1(a) [ŋ:g] is a prestigious MC form. The social patterning is inconclusive for 1(a,b) and 2(b), possibly because (ng - 1) is a regional norm and in conflict with the national norm (ng - 2). The local norm is defended by such arguments as that it must be correct to pronounce one's /g/s properly.

## 3 *Variation and performance*

Linguists like to describe language in terms of stable systems and categorical rules. Variation is something of an embarrassment which has to be accounted for. This point has been argued and demonstrated convincingly by socio-linguists on several occasions (see e.g. Labov 1969, 1972; Bickerton 1971; Trudgill 1974). What has been less clearly perceived is that variation also calls for a special kind of phonetic theory.

Phoneticians have, for obvious reasons, concentrated on the ways in which the vocal apparatus brings about phonological contrasts. The speech signal does of course contain much more than strictly linguistic information. For instance, a doctor listening to a patient's voice for signs of disease is looking for personal information, and will have little use for linguistic information. Now it is clear from the study of Scouse shibboleths that socio-linguistic information partly overlaps with linguistic information and is partly something different. This is why variable data cannot be adequately analysed by means of linguistic phonetics alone.

What we need is some kind of performance model. To set up a model of speech performance is beyond the scope of this paper (see Fromkin 1968, 1971; Laver, 1971). On the other hand, variation itself provides evidence for what such a model must include. What we can do is to construct a partial and crude model of those parts of the process which are subject to variation. The model given in Figure 1 relates segments to their phonetic realizations.

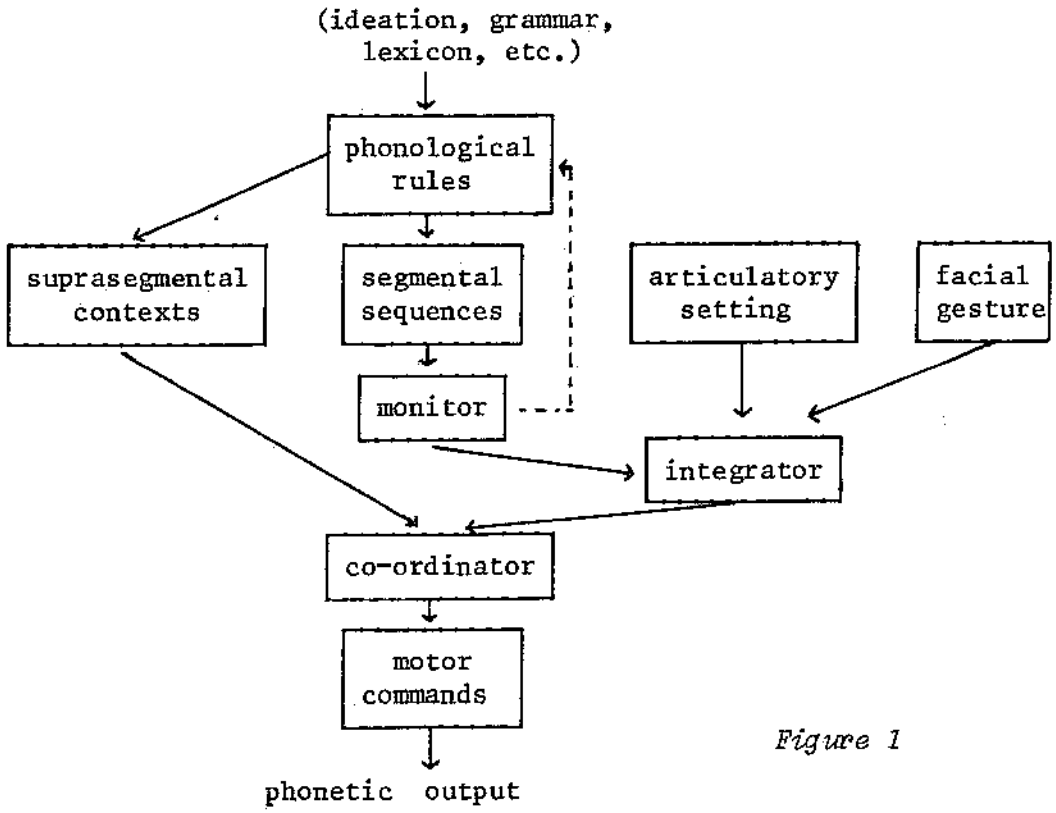



*Figure 1*

We are not concerned with the storage or retrieval of linguistic information, nor with the way linguistic units are selected to convey the speaker's message. We start with his message as a string of segments in its suprasegmental context. The notion of 'segment' is deliberately left vague, but is more like an extrinsic allophone (as defined by Ladefoged 1965) than the classical phoneme. Speech production is constantly being monitored, and some outputs at this stage are rejected and replaced. Successful outputs are going to be converted into muscular movements, and these have to be carefully integrated with other movements -- arising e.g. from articulatory setting and facial gesture -- which are not phonological in origin. For example, a smile will interfere with lip rounding for a vowel, and the speaker has somehow to reconcile contradictory instructions to the orbicularis muscle. The instructions for the different parts of the message have to be very carefully co-ordinated in time to produce the motor commands for well-formed speech.

Each box, its position relative to other boxes, and its major functions, is probably universal and belongs to general phonetic theory. I shall argue here that the detailed working of each box is not fixed but variable, and that this variation is the origin of many low level rules of considerable socio-linguistic significance.

3.1 *Strict phonological variables* While much of the continuous variation in speech can be explained by our model, there are a number of variables which originate at a much deeper level than phonetic realization. They involve differences of several kinds:

(i) *Inventory*, e.g. Scouse has five short vowels corresponding to the six of RP, so that words having /ʌ/ in RP have /ɷ/ in Scouse.
(ii) *Phonetic targets*, e.g. Scouse [ɜ:, ë:, ɵ:] for /ɜ/ cannot be explained by low level rules, and so probably start off as different 'targets'.
(iii) *Lexical incidence*, e.g. Scouse /luk, kuk/ 'look, cook' corresponding to RP /lɷk, kɷk/, where both Scouse and RP have both vowels /u/ and /ɷ/ but distribute them differently in the lexicon.
(iv) *Combinations* of the above, e.g. Scouse /wɒn/ 'one' corresponding to RP /wʌn/ combines a difference of incidence with a difference of inventory.

Variables of these kinds are of course perfectly familiar and well understood.

3.2 *The Monitor* We suggested above that some variables involve not so much different targets as the modification of the speaker's natural target. The 'monitor' is the box that carries out this modification. In fact, monitoring probably goes on at every stage of speech production: e.g. we might avoid swear words in front of vicars, or difficult words and syntax when talking to children, or we might consciously reject several forms of a sentence when writing prose before getting it right. Only the segmental monitor is included in the model because this is the only one for which I found interesting data in Scouse of general socio-linguistic significance.

Much of the inconsistency in middle class speech might be explained as the result of competition between dialects, so that sometimes the Scouse form wins, and sometimes the RP form. (A compromise form is then presumably a draw.) However, competition cannot explain such hypercorrections as /bʌtʃə/ 'butcher', /gɑ:s/ 'gas' or /wɛəm/ 'worm'.

An alternative explanation is that the speaker encodes words in a single segmental form, and monitors them before they are uttered, and 'corrects' some of them. To switch from Scouse to RP, the speaker has to know (1) how to modify a particular segment, and (2) what words it is appropriate to modify it in. Imperfect modification leads to compromise forms and modification in inappropriate words gives rise to hypercorrection.

An informal study of my own slips of the tongue suggests that there are in fact two kinds of hypercorrection. The first and familiar kind involves individual lexical items. I get easily confused over words like *putty* and *nasty*, because I am not sure whether an RP speaker says /pɒtɪ/ or /pʌtɪ/, /næstɪ/ or /nɑstɪ/; and I have to look them up in a pronouncing dictionary. Other words with RP /ʌ/ or /ɑ/ are perfectly all right, as long as only one of them occurs in a tone-unit and there are no other words with /ɷ/ or /a,æ/. Compare the hypercorrections:

(2) I've just pulled /pʌld/ a button /bɒtn/ off my jacket.
(3) Good luck! /gʌd lɷk/

Now I know perfectly well how RP speakers pronounce *pull*, *button*, *good* and *luck*, and I would get them right in isolation. The second /ɷ/ triggers off the modification rule, but it is the first /ɷ/ that gets modified. *A good book* is a difficult phrase because I feel I have missed an RP /ʌ/; *cup-hook*

pronounced /kʌphɔk/ is very suspicious, but actually right. Comparable mistakes with the variable (ɑ) include:

(4) and the wolf ran /rɑn/ after /æftər/ ɪt.
(5) the black castle -- ðə blɑ:k kæsl.

In the case of (ə) I have stopped myself saying things like /kɛəli hə/ 'curly hair', but have not collected any slips actually uttered. This hypercorrection seems to occur within the tone-unit, as in this example:

(6) it's ˌ not a jug /dʒʌg/, but it's as ˈgood /gʌd/ as a jug /dʒɔg/.

I know how RP speakers pronounce *jug*, and accordingly modify my /dʒɔg/ in the first tone-unit; in the second tone-unit, the modification was shifted on to the earlier /ɔ/ of *good*, turning it into /gʌd/.

These examples seem to confirm Laver's speculation (1970) that speech is not produced segment by segment, or syllable by syllable, but in short stretches of about the size of a tone-unit. The monitor does not hold up speech production while it rejects a segment and finds a substitute; it sees the segment coming and changes it on the production line. The hypercorrections can be explained if the objectionable segment is foreseen not in its segmental context or syllable, but just in its tone-unit; the monitor sets up the modification rule and applies it to the first segment in the tone-unit that comes along the line and happens to fit.

3.3 *Articulatory setting* Because phoneticians have concentrated on linguistic phonetics, we describe articulation with reference to the minimum number of organs which are necessarily involved. When describing /b/, we refer to the lips, the velum and the vocal folds, but we do not specify the width of the pharynx, or ask what the mandible is doing. Much of this secondary activity facilitates the primary articulation -- e.g. efficient voicing of stops depends on a sufficiently low supra-glottal air pressure, which in turn depends on the state of the pharynx -- but it is also socio-linguistically determined. For any dialect, the organs of the vocal tract have certain preferred positions, which may be different from those they have in the physiological state of rest. The movements away from and back into the preferred positions give a characteristic 'colour' to that variety of speech. The preferred shape of the vocal tract is known as the *articulatory setting*.

The Scouse setting has been described in detail elsewhere (Knowles 1974a). The parts most relevant here are the setting of the velo-pharyngeal mechanism, and the setting of the jaw. For the former, the centre of gravity of the tongue is brought backwards and upwards, the pillars of the fauces are narrowed, the pharynx is tightened and the larynx displaced upwards. (If this sounds complicated, it is actually rather like the first stage of a swallow!) The lower jaw is typically held close to the upper jaw, and this position is maintained even for 'open' vowels. The main auditory effect of this setting is the 'adenoidal' quality which is so characteristic of Scouse and which is produced even though the speaker's nasal passages are quite unobstructed. An important result for articulation is that the tongue moves in a restricted space, and the tip loses much of its normal mobility.

3.4 *Gesture* Gesture is peripheral to our model because it is by definition paralinguistic and is primarily visual anyway. It is relevant in so far as it competes with setting and phonology to determine the movements of the face. In addition, a regionally restricted gesture might produce unfamiliar auditory results which are difficult to distinguish from other kinds of auditory information. For example the Scouse 'hostile look' involves spreading the upper lip and pushing it upwards and outwards to reveal the upper teeth; this turns lip rounding into labio-dental approximation, with marked auditory results. If a Scouser exclaims 'Ooh!' with this gesture, it sounds like an extended syllabic variety of /r/; in order to explain such an oddity we have to take gesture into account.

3.5 *Integration* In some cases, the movements needed to realise a segment are compatible with setting and gesture, so that integration is relatively simple. Take Scouse [t] for instance. The back of the tongue is raised, and in view of the restricted space and immobility of the tip, the alveolar contact is most easily made with the blade: Scouse [t] is thus velarized and laminal.

In other cases, conflicting demands have to be reconciled. To produce an [s], the tongue is grooved to direct a narrow jet of air against the cutting edge of the teeth. It is easy enough to make a velarized laminal [s], but the tongue shape makes it difficult to groove it efficiently; this allows air to escape more diffusely and the consonant sounds more like [ʃ]. The speaker can give priority to the phonology, override the setting and produce the consonant efficiently; or else he can give priority to the setting and accept the consequent distortion of the consonant. Both types are heard in Liverpool; priority to the setting makes for a 'thick' accent and seems to be common among younger speakers.

Vowels -- other than close back ones -- are incompatible with the setting, at least as vowel production is traditionally understood. What is understood as vowel quality depends not so much on a particular articulatory position as on the auditory correlate of the lower formants. Corresponding vowels in Scouse and RP are recognizably similar in quality, but whether they are produced with anything like the same shapes is an open question. Tongue shape must be affected when an 'open' vowel is produced with a relatively close jaw, or a 'rounded' vowel produced with spread lips. (In the latter case, 'rounding' can be achieved by raising the tongue tip slightly, but not enough to produce r-colouring.)

Now it might at first seem sufficient to plot variations in vowel sound along the traditional dimensions of height, retraction and rounding. However, an RP speaker can produce his /ɛə/ so that it comes well within the limits of a Scouse /ɜ/ without for a moment sounding like a Scouser; conversely, a Scouser can modify his /ɜ/ to [ɛə] without sounding remotely like an RP speaker. Socio-linguistic information in vowel sounds only partly overlaps with vowel quality as traditionally defined. There is an important area here of articulatory and acoustic phonetics which has yet to be investigated properly.

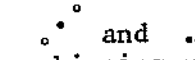
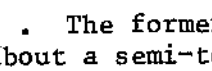
The integration of phonology and setting is complicated for consonants and vowels by the third factor, namely gesture. Scouse tends to give priority to gesture. A grin spreads the lips, even for a 'rounded' vowel, and despite the preferred neutral lip shape. A gawp opens the jaw, even

for 'close' vowels, and despite the close jaw setting.

It is said of both Scousers and RP-speakers (and Englishmen in general for that matter) that they speak without moving their mouths. While this may make intuitively good sense, as a matter of observation Englishmen move their mouths constantly, and Scousers look very different from RP-speakers. The comment can perhaps be interpreted in terms of the priority given to the optimal realization of phonological segments. Phonological movements in Scouse are greatly affected by setting and gesture. RP treats vowels and consonants differently: vowels seem to have a low priority, whereas consonant articulations -- especially rounding for /r ʃ ʒ tʃ dʒ/ -- may be made vigorously and override the preferred setting.

3.6 *Suprasegmentals* Segments, integrated with setting and gesture, are grouped into syllables, and syllables are grouped according to the positions of the stresses in the tone-unit. The rules involved are peculiar to English, and are different in other languages; but there is probably not much variation inside English which is of interest here. The most important stresses, however, are realized by a pitch movement or *tone*, and this is an important source of variation.

In the first place, dialects may have different inventories of tones. Scouse makes a distinction not found in RP between tones which glide smoothly up or down, and those which skip suddenly between one pitch level and another. Thus if *yes* is said on a rising pitch, there is a contrast between

 and  . The former is similar to the RP low rise. The latter, skipping up about a semi-tone, is common in Scots and Ulster English, but is quite foreign to most dialects of England. For yes/no questions, Scousers use a tone which skips from high to low pitch, and this has no counterpart in RP.

Dialects also differ in the uses to which tones are put. An RP low rise on a greeting like *Morning* corresponds to a rise in Scouse, but where an RP-speaker uses the rise in a list, the Scouser might use a series of tones skipping down. The RP fall-rise for a transferred negative corresponds to the fall-rise in Scouse, but in other contexts Scouse will use a skip down or even a rise-plus-fall. In common with other Celtic dialects, Scouse uses a nuclear rise-fall where RP has the simple fall.

3.7 *Co-ordination* The co-ordination of different kinds of speech activity has not figured largely in studies of variation, but is in fact the source of a considerable amount of socio-linguistic information. This information is linguistically redundant, and so is not dealt with in traditional phonetics; nor can it be transcribed very easily. Consequently, the analyst may perceive it well enough, but there is not much he can do about it. For convenience, we shall distinguish three kinds of co-ordination: (1) segments and the breathing cycle, (2) segments and suprasegmentals, and (3) segments and other segments in sequence. Co-ordination is essentially concerned with the relative timing of speech events.

3.7.1 *The breathing cycle* When we speak, we use the lungs as a kind of bellows for the production of segments, but we still have to use them in their primary function of respiration. The time ratio of inspiration to expiration is normally about 1:1, but in speech the time of inspi-

ration may be more than halved, and that of expiration doubled. Expiration has to be carefully controlled to maintain a roughly level sub-glottal air pressure; the rate of expiration depends on muscular action and the obstruction to the air-flow at the glottis.

Although it is difficult to describe the difference exactly, Scousers seem to control expiration differently from RP speakers. I seem to use up air faster when speaking Scouse than in my Merseyside modified RP. Many Scousers use a breathy voice quality, which allows a rapid air-flow, and together with the velo-pharyngeal setting, gives the impression of congestion in the nasal passages.

Increases in the air-pressure in exhalation are involved in the realization of stress. The rate of increase is another possible variable, being greater with the tense articulation of *staccato* speech than with the lax articulation of *legato* speech (for these terms, see Crystal 1969:164-5). In the British tradition of stress marking (Kingdon 1958; O'Connor and Arnold 1973), a distinction is made between major stresses which have pitch prominence, and minor stresses which do not. This works well enough for RP but not for Scouse. In Scouse, it can be very difficult to decide whether a syllable is supposed to be stressed or not, unless it has pitch prominence; variations in loudness and duration are insufficient on their own.

The switch from ordinary to speech breathing seems to be made by an intake of air, irrespective of the stage at which it occurs in the breathing cycle. There is little scope for variation here. The switch back from speech breathing to ordinary breathing is unremarkable in RP, whereas in Scouse the rapid exhalation of ordinary breathing often begins just before the end of the utterance. This means that final vowels end voiceless or in [h], e.g. [poh, ʃʉoh] 'poor, shoe'; final consonants are aspirated, e.g. [hamʰ, pʰɒbʰ, bɒsʰ] 'ham, pub, bus', and voiceless consonants may even be pre-aspirated, e.g. [tʰahkʰ, bɒɒhtʰ] 'tack, boat'. It is this rush of air that makes audible the release of a final nasal, giving the impression of a weak, devoiced oral stop, as in [wɒnd̥ʰ] 'one'.

3.7.2 *Suprasegmentals* The British concept of a 'tone' implicitly recognizes that the way segments fit into a pitch pattern may be linguistically significant. For example, the *fall* and *rise-fall* are treated as different tones, even though in many contexts they involve very similar pitch patterns:

(i) *fall* ˚ · · / təmɒrəɒ (ii) *rise-fall* · ˚ · / təmɒrəɒ

In *tomorrow*, both tones involve rising-falling pitch; the tone in each case is identified as that part of the pitch pattern which follows the peak of the stressed syllable.

The details of the fit between segments and pitch movements can be of socio-linguistic significance. In the RP pronunciation of *to'morrow*, the /m/ is already at about the pitch level of the beginning of the fall; in this way the tonic fall stands out against the low pre-head on the first syllable. In Scouse, the maximum pitch is more likely to be delayed until the peak of the stressed syllable, so that stress is marked by a change of pitch direction; this means that the /m/ and even the beginning of the vowel are on the

preliminary up-glide. (For this reason the Scouse fall is easily confused with the rise-fall.) There are in fact about half a dozen tones in Scouse which fit segments differently on to a rising-falling pattern, and make the melody of Scouse so different from RP and other varieties of English.

Dialects also differ in what happens to the pitch pattern after the syllabic peak. Variants of the rise-fall are in the handbooks (e.g. O' Connor and Arnold 1973:9). The RP fall is generally described as being smartly completed on the stressed syllable, leaving any following syllables on a low pitch. Scouse tends to have a much more general fall; in *to'morrow* the fall beginning on /ɒ/ gradually falls over the weak syllable /rəʊ/.

A rather different problem of segmental fit concerns what might be called dialect 'rhythm'. The term is being used here in a very special sense (Knowles 1974b) and refers to the way the time of the utterance is distributed among the segments. The duration of vowels, for instance, is influenced by a number of factors:

(i) Phonological length: long or short.
(ii) Degree of openness.
(iii) The nature of the following consonant, if any.
(iv) Tempo.
(v) Its position relative to stress.
(vi) Its position in the tone-unit, before or after the nucleus.

Each factor influences duration in any kind of English, but their relative influence, particularly in the case of the first three, is an important variable. In Northern English, for instance, vowel duration depends largely on phonological length, variation according to openness being minimal, and the effect of the following consonant being almost imperceptible. Thus the vowel of *leek* is fully long, even though it is close and followed by a voiceless stop; the vowel of *jam* is very short, even though it is open and followed by a voiced consonant. In the South of England, on the other hand, openness and the effects of the following consonant can completely distort differences of phonological length, so that the 'short' vowel of *live* can be longer than the 'long' vowel of *leek*, and the 'short' vowel of *jam* can be very long indeed. Scouse rhythm is different from both the Northern and Southern types, but shares some of the features of both. It has the Northern short vowel in *jam*, and Southern vowel shortening in *leek*; the vowel of *leg* may be lengthened in a way foreign to the rest of England, but quite common in parts of Scotland and Ireland.

3.7.3 *Segment sequences* In phonology, segments can be treated as items in a string; phonetically, there is considerable overlap in the production of neighbouring sounds. Thus in a word like *square*, one of the first articulatory movements is likely to be the lips moving into position for /w/, even though this is ordered phonologically after /s/ and /k/. Consequently, if a segment is analyzed as a matrix of features, the realization of these features is staggered in time, some being prepared before others.

There is of course enormous scope for variation in the overlap between segments. There are three kinds of variation needing separate treatment: (1) preparation for articulation, (2) phonation, and (3) the realization of sequences of vowel segments.

3.7.3.1 *Preparation* If an organ is not involved in a given articulation, it has still got to be doing something, and subject to the demands of setting and gesture, it can either prepare for a future articulation or remain in the position for a preceding one. Anticipatory and perseverative assimilation are to be expected in normal utterance. Indeed, it is often the absence of assimilation that is significant and noteworthy, rather than its presence. A remarkable feature of Scouse is the frequent failure to prepare for an articulation.

In syllable initial clusters like /bl, kw, tr/, one can expect the second consonant ($C_2$) to be prepared during the articulation of the first ($C_1$), and this can affect the quality of $C_1$, so that, for instance, /t/ is retracted before a following /r/. In Scouse the preparation may be delayed, so that the organs are not fully ready for $C_2$ by the time $C_1$ is completed. The awkward transition from $C_1$ to $C_2$ gives the odd impression that the two consonants belong to different syllables. In extreme cases there is even a slight gap filled with a nondescript vowel element, thus [bᵊlak, gᵊrɛɪ tˢᵊwɛntl, ʃᵊrɪŋk] 'black, grey, twenty, shrink'. This epenthetic [ᵊ] is most prominent when the syllable it occurs in has a nuclear tone, since by the rules of Scouse intonation, its pitch moves in the opposite direction to the tone proper.

The reduced preparation time has a further effect in the case of /r/. In a word like *grey* in RP, there is ample time for the tongue and lip positions for /r/ to be prepared, and for the articulation to be carried out under careful control. But in Scouse, with the delayed preparation, there is insufficient time to carry out a controlled approximant; instead, /r/ becomes a rapidly executed tap against the alveolar ridge.

In the case of CV sequences, the target vowel quality may not be reached by the time the consonant is released, and a glide may be heard from a close central [ɨ] quality to the close vowels /i, u/. Thus [θrᵗɨ, tˢᵗu] 'three, two' and [bᵗɨn, bᵗʉk, skᵗʉl] 'been, book, school'. These sequences of glide plus vowel are re-interpreted as diphthongs, and since Scouse diphthongs are typically of the 'falling' type, the stress is shifted on to the first element, hence the alternative forms [θŕɨ, tˢ́o, b́ɪn, b́ok skóol].

Delayed preparation leads to the fracture of a vowel followed by a consonant made with the tongue. In *queen*, the tongue may produce an audible intermediate [ə] quality as it moves from the front vowel to the consonant. This also happens in diphthongs, e.g. [steɪᵊl] 'stale', because in this dialect the second elements of diphthongs are not weakened.

3.7.3.2 *Phonation* Movements of the larynx have to be co-ordinated with articulation to produce voicing at the right time for voiced segments, and the way this is done varies from dialect to dialect. In a widespread Northern type, voicing starts close to the beginning of an initial voiced segment, as in [du] 'do', and very closely after the end of a voiceless segment, so that a voiceless stop as in [tu] is perceived as unaspirated. In the RP type, the onset of voicing is delayed, so that initial /d/ is devoiced, and /t/ aspirated. In final position, the Northern type continues voicing to the end of a voiced consonant, whereas the RP type cuts voicing off early, so devoicing a final voiced consonant. Scouse agrees with the RP type rather than the Northern type, and goes even further in pre-aspirating voiceless consonants as the result of the utterance final breath rush.

3.7.3.3 *Vowel sequences* The vowel is the centre of prominence in the syllable, and so if two vowel elements occur in the same syllable, the question must be raised how prominence is distributed between them. The second question, whether the vowels in sequence belong to the same syllable or not, is how the speaker moves from one to the next.

English diphthongs characteristically take the stress on the first element. There are a few exceptions, like /ju/ of *view*--- which is made to conform to the falling type in some Welsh dialects, hence [vi̯o] -- or /jɔ/ in *pure* or the /j / of *here* in one kind of RP. Otherwise there is little variation. Where there is considerable variation is in the distribution of prominence between the two elements.

The commonest English type, which can actually be traced back to Germanic, concentrates prominence on the first element. Any extra duration is shifted to the first element, thus [lɑɪt, lɑ:ɪd] 'light, lied'; the quality of the first element is produced clearly, and the transition away from it to the second element is very slow, and the 'target' quality of the latter may not be reached at all. Scouse diphthongs vary between this and a 'Celtic' type (discussed in Knowles 1974b). In certain environments (the details of which need not be discussed here), it is the second element that attracts the extra duration, thus [braʊ:n, nɔʊ:z] 'brown, nose'. In addition the transition from the first element to the second is very rapid, and the quality of the second element is made clearly, while that of the first may be obscured, thus [həʊs, nəɪt] 'house, night'. The second element is thus much more prominent than in RP, and this effect is even greater when the diphthong occurs on a 'skipping' tone, because the skip can actually occur between the two elements. Thus in [həʊs], the [ʊ] might be a semitone higher or lower than [ə], and the general impression is that the two elements belong to different syllables.

The different treatment of the first element leads to possibilities in RP which are not available in Scouse. Close vowels are made intrinsically more prominent by making them more open, and RP /i/ in *we* can become the more open [ɪ] in [wɪə] 'we are'. This [ɪ] can be made more open still, so that *we are* becomes indistinguishable from *where*. In the isolated case of *really*, opening the first element makes the word similar to *rarely* for some RP speakers. The change of /ʊə/ to [ɔ:] is a parallel change at the back. Scouse does of course have [ʃɔ:] for *sure*, and these forms are more likely to have been introduced ready-made from outside, than to have developed by the native rules of Scouse.

If two vowels belong to different syllables in RP, close vowels can be laxed, as in [ðɪ ɪst, tʊ ɪt] 'the East, to eat', and other vowels can be separated by a linking or intrusive *r* (on *r* see J. Windsor Lewis 1975). Other devices are used sporadically, such as the glottal stop. In Scouse, by contrast, close vowels are followed by homorganic glides, thus [ðiʸɪst, tuʷɪt], and the first vowel can be reduced almost to [ðəʸɪst, təʷɪt]. Other vowels are followed by [ɾ], which is executed rapidly and may become a flap, thus [ɹɔ:ɾɛgz] 'raw eggs'. In the case of reducible grammatical words, there is some variation in the ordering of glide insertion and vowel reduction. Normally, the glide is inserted and then the vowel is reduced, but very occasionally reduction applies first, giving such examples as [ðəɾɛkɔʊ, jəɾɑ:nt] 'the Echo, you are not'. In these cases, the vowel has been reduced to schwa, so that the glide inserted is regularly of the *r*-type.

In sequences of three vowel elements, RP maintains the prominence of the first and weakens the second. The quality of the second is unlikely to be reached, and it may disappear altogether, thus the monosyllable [tɑ:ə] 'tower' and the disyllable [gə:ɪŋ] 'going'. In Scouse the first element may be obscured, and the second serves as a glide to the third, thus [tawə, təwə] 'tower', and [gɔwən] 'going'.

3.8 *Motor commands* The motor commands for speech depend on the working of the various boxes we have investigated which process the output of the phonological rules. The motor commands themselves must vary within limits. It is a commonplace of phonetics that it is impossible to produce exactly the same sound twice. The speech apparatus has mechanical and neuro-physiological limitations; for example, it is difficult to control a trill to a given number of taps, and impossible to do so to less than a whole number. (One cannot produce a trill with two and a half, or three and three quarters taps.) In real speech, variation is acceptable over a much greater area than the natural limits, and this area is subject to socio-linguistic control.

Consider for example the control exercised in the production of a stop. In the 'cardinal' kind of stop, there are three stages: a complete closing off of the air-stream, holding the closure, and finally a smart release. As the stop is released, there is necessarily a short period of affrication, but the plosion is so controlled that the friction is perceived as part of the stop and not a separate element in itself. Stops in RP are generally close to this cardinal type. Scouse has a different rule: for stops at the beginning of a word or a stressed syllable, the articulators close off the air-stream. The release varies freely, and can be smart as in RP, or very gradual, producing considerable audible friction as in [tˢʰɛn] 'ten'. In other positions, where it is not specified by the rule, there may not even be a proper closure. The articulators move into close approximation and then out again, thus [dɔ:ˢɛ dɒˣɛ ɷΦ] 'daughter, docker, up'. These close approximants -- marked with the raised symbol -- do not sound like ordinary fricatives. First, they are encoded as stops, and affect the duration of the preceding vowel as stops rather than as fricatives. Second, they do not reach a steady state, but are constantly varying in the degree of approximation. These incomplete Scouse stops have sometimes been described to me as 'affricates', but this is not accurate at all: they do not fit into any of the familiar phonetic categories.

Perhaps we can include in this section a note on rounding. As a phonetic category it might seem simple enough, but there is enormous variety in the way it is produced, and different methods have slightly different auditory results. The type that springs to mind is the one a phonetician might use to produce Cardinal 8, namely a vigorous protrusion of the lips to make a small ring-shaped aperture. A different type often seen in Belfast has protrusion with a large aperture which exposes the front teeth. In the West Riding, rounding is sometimes accompanied by a forward movement of the whole lower jaw. Several kinds of rounding are found amongst RP speakers, but they generally involve very little protrusion or lip-rounding. In Scouse, for reasons given above, rounding may be produced by the tongue, or by a movement of the lower lip towards the upper teeth; where setting and gesture permit, rounding is often made by a vertical approximation of the lips at the corners of the mouth.

3.9 *Variation and individuals* We have so far traced a number of different kinds of variable to their source in speech production. We must now explain variation in the speech community.

Variation among individuals, or groups of individuals, is fairly straightforward. People have different grammars, and slightly different programmes for the realization of phonological sequences. An individual changing from one style to another makes changes in his grammar, and also in some or all of the boxes in the realization programme. Stylistic changes include those which distinguish one social group from another, but also include speaking with a 'slight' or 'thick' accent, or speaking 'clearly' or 'indistinctly'. A 'thick' accent is not the same as WC speech, and educated MC people are quite capable of mumbling. A number of MC Liverpudlians will adopt something close to RP phonology, and retain a thick accent in the way they realize the segments, and conversely, some WC people have Scouse segments, but have only slight accents, on account of their realization rules. (In this connection, it is interesting to observe the Mancunian Brummie and other weird forms of speech which are accepted as Scouse on stage. An actor may imitate odd bits of the phonology and the realization, which is often manifested in a cold-in-the-head voice and a funny intonation.) The stylistic variation which an individual is capable of seems to depend on what he himself has learned to vary; *contra* Trudgill (1974:134) I see no reason to suppose that speakers internalise a diasystem. People have remarkably little idea how others speak.

A full account of variation will explain how it comes about in the first place, and how it can lead to sound change. A distinction must be made between variables which are produced natively in a dialect, and those which are brought in from outside. Many variables in Scouse are taken from RP, and their spread is a matter of fashion, along with rotary driers and platform shoes, and it is therefore not a linguistic problem. Native variables probably derive from a re-interpretation of the information in the speech event. The Scouse setting, for example, may well have begun in the speech of some individuals as personal information, and was then taken as socio-linguistic information, and adopted by the whole community. The first element of the vowel [ɵo] probably began as a glide to the vowel proper as the result of delay in preparation; it was then re-interpreted as the first element of a diphthong.

Rather different questions are posed by variation in the speech of an individual in a single style. Most of the variation reflects the limitations of the speaking process -- just as, say, variation in sliced loaves reflects the limitations of the baking process -- rather than any inconsistency on the part of the speaker. It is very easy for the student of variation to exaggerate the amount of inconsistency; in fact, inconsistencies are just intermittent oddities in the stream of speech, and it would be very surprising if this were not the case.

Real inconsistencies in Liverpool speech tend to involve the shibboleths described above. For instance, a MC speaker may sometimes use /ʌ/ where RP has /ʌ/, and sometimes /o/. The fact that inconsistency occurs at all could be explained as a random choice by the speaker between competing forms; but we must also explain why the prestige form is selected more often in formal styles and by higher social groups. The monitoring process, which

ensures that only the right segments are passed through for realization, must be variable in the efficiency of its operation. It might be asked exactly how the efficiency is changed, but this is not a question specific to phonetics or socio-linguistics: there are many spheres of human activity where the efficiency of quality control in a factory varies according to the quality of the goods being produced. An alternative explanation (Trudgill 1973:156) is that the speaker has actually internalized some quantitative knowledge, but *pace* Trudgill I see little connection between phonological variables and stylistic repetition and the use of swear-words. Variable percentages are surely a measure of performance, and that is no reason to ascribe them to competence.

## 4 *Variation and transcription*

It is customary in linguistics to treat phonetic transcriptions as raw data: the transcriber encodes a speech event as a sequence of symbols, and merely converts auditory information into visual information. This is in general a perfectly valid procedure, because transcription has been developed precisely to extract linguistic information. There is no reason to assume in advance that socio-linguistic information can be treated in the same way.

In phonetics -- as opposed to linguistics -- transcription is a useful tool for everyday use, but it is not designed to symbolize all the patterns explained by phonetic theory. For example, the state of the glottis and voice onset time are two continuous and separate scales, but they can be transcribed only clumsily by the use of discrete symbols and diacritics, e.g. [t͡ʔ tʰ t̬ ɖ d]. Phonological transcription presents no problem because the symbols used represent conventional ranges on one or more scales. A phonetic transcription, on the other hand, merely represents those pieces of information which the transcriber happens to notice, and for which he has a symbol or diacritic. It would be extremely naive to suppose that an impressionistic transcription, however 'narrow' or apparently 'accurate' in any way uniquely represents the original speech event.

When we transcribe variables, we perceive socio-linguistic information of several kinds, and we represent it by our symbols and diacritics with their conventional values. Segments are as it were 'pigeon-holed' on a vowel or consonant chart. The easiest fallacy of all is to assume that the original sound has the same properties as the pigeon-hole it is put into, e.g. if some occurrences of /t/ are transcribed [t], some [ts] and some [s], then the variable must involve the scale stop versus affricate versus fricative. To avoid this fallacy, the socio-linguist has first to understand the variation in general phonetic terms, and then choose and re-define his symbols precisely and *ad hoc*.

In traditional dialectology, astonishing virtuosity is sometimes found in the transcription of dialect forms. There are just two problems. First, the detail goes far beyond what any phonetician could give with any confidence, accuracy or consistency. Second, unless the reader is familiar with the dialect in question, the detail still does not suggest the original form. One has to imagine a 'Yorkshire voice' in order to read a transcription of Yorkshire dialect, a 'Texas voice' to read a Texan transcription,

and so on. In other words, one has to provide much of the very socio-linguistic information which the transcriber set out to represent in the first place.

Impressionistic virtuosity can now be replaced by mechanical accuracy, as when Labov (1972) uses the spectrograph to make accurate measurements of vowel height for his (eh) variable. Measurements do not in themselves prove a theory -- otherwise there would be few controversies in natural science -- and measuring F1 does not prove the correctness of plotting (eh) on the scale of vowel height. It demonstrates that vowel height is involved somewhere, but so, possibly, are duration and diphthongization, and these can also be measured on spectrograms. (In linguistic phonetics, one might measure the duration of English /t/ and /d/ in certain environments, and find that /t/ is longer, but it would be extremely naive to assume without further experiment that consonant length is therefore a distinctive feature in English!)

These basic phonetic problems in variation have not, as far as I know, been even recognized by socio-linguists, and it may seem mere quibbling to bring them up now. They were very real practical problems in the study of Scouse. I spent several months transcribing my material, and working out correlations between variables and social groups, and it was clear I could make a convincing-looking case, complete with diagrams, tables and graphs. However, I became increasingly aware of the inaccuracy of the transcriptions, and the inadequacy of the variable scales, and this led me eventually to throw away my notebooks and start all over again.

Published statistics of linguistic variation are very impressive from a sociological point of view, but they are based on very questionable foundations. The data are much more difficult to quantify than in some other kinds of survey, with questions like 'Do you vote Labour?' or 'When did you buy your last toothbrush?'; we should have little confidence in an opinion poll which could not definitely distinguish the reply 'I vote Labour' from 'I vote Conservative'. Until socio-linguists improve their technique for identifying and quantifying phonetic forms, it must remain in doubt exactly what their statistics are proving.

## 5 *Conclusion*

There can be no serious doubt about the relationship between variables, style, and social group. Variation in speech is not free, but is governed by the speaker's social standing, and by the social situation in which he is speaking. Indeed, these social influences are so powerful that they readily show up in any survey despite the crudity of the investigative techniques that might be adopted.

In many cases, crude techniques are all that are required or available. For example, in a study of social levels in the Arabic of Baghdad, there would be plenty of material in the gross differences, without going into phonetic minutiae. Secondly, an excellent way of teaching phonetics or socio-linguistics is to set a simple project involving variation, and gross differences are enough for the beginner to deal with.

In his early work on New York English, Labov was obviously right to treat variables as he did, in order to demonstrate the connection with social forces. It does not in any way belittle his achievement to point out that in future work variation might be looked at more precisely. Nor is there any necessary conflict between Labov's view of variables and the one outlined here. For a sociological study of the speakers of a dialect, gross variables are enough; but more detail is required for a phonological description of the dialect itself.

The study of urban dialects in the United Kingdom has scarcely begun, and yet the speech of the vast majority of people is influenced by a small number of conurbations including London, Birmingham, Liverpool, Manchester, Leeds-Bradford, Newcastle, Glasgow and Belfast. For WC speech, it would be quite easy to find characteristics which distinguish one conurbation from another. It would be impossible to study MC speech without going into variation, and without going into the kind of detail outlined here. Despite intensive study of English phonetics, very little is known beyond the professional opinions of phoneticians of exactly what kind of pronunciations are generally received among educated and middle class people.

General phonetic theory has concentrated on linguistic information, the mechanisms which produce it, its acoustic description, and the way it is perceived. The task facing the urban dialectologist is to do the same thing for socio-linguistic information. This means treating variation as part of performance rather than competence, and tracing it to its origin in speech production.

## REFERENCES


Bailey, C-J.N. and R.W.Shuy (1973 eds) *New Ways of Analyzing Variation in English.* Georgetown UP.

Bickerton, D. (1971) 'Inherent variability and variable rules', *Foundations of Language* 7:457-92.

Crystal, D. (1969) *Prosodic Systems and Intonation in English.* Cambridge, CUP.

Fromkin, V. (1968) 'Speculations on performance models', *JL* 4:47-68.

(1971) 'The non-anomalous nature of anomalous utterances', *Language* 47:27-52.

(ed 1973) *Speech Errors as Linguistic Evidence.* The Hague, Mouton.

Kingdon, R. (1958) *The Groundwork of English Intonation.* London, Longman.

Knowles, G. (1974a) *Scouse: The Urban Dialect of Liverpool.* Unpubl. PhD thesis, University of Leeds.

(1974b) 'The rhythm of English syllables', *Lingua.*

Labov, W. (1966) *The Social Stratification of English in New York City.* Washington D.C., Center for Applied Linguistics.

(1969) 'Contraction, deletion and inherent variability of the English copula', *Language* 45:715-62.

(1972) 'The internal evolution of linguistic rules', in Stockwell and Macauley eds.

Ladefoged, P. (1965) 'The nature of general phonetic theories', *Georgetown Monograph Series on Languages and Linguistics* 18.

Laver, J. (1970) 'The production of speech', in Lyons (ed. 1970).

(1973) 'The detection and correction of slips of the tongue', in Fromkin (ed. 1973).

Lyons, J. (ed 1970) *New Horizons in Linguistics*. London, Penguin.

O'Connor, J.D. and G.F.Arnold (1973) *Intonation of Colloquial English*. London, Longman.

Stockwell, R.P. and R.K.S. Macauley (1972) *Linguistic Change and Generative Theory*. Bloomington, Indiana UP.

Trudgill, P.J. (1973) 'Phonological rules and sociolinguistic variation in Norwich English', in Bailey and Shuy (eds 1973).

(1974) *The Social Differentiation of English in Norwich*. Cambridge, CUP.

Windsor Lewis, J. (1975) 'Linking /r/ in the general British pronunciation of English', *JIPA* 5:37-42.