Precision

Mark Tatham
Katherine Morton

Reproduced from *Occasional Papers* 23, Department of Language and Linguistics, University of Essex, Colchester UK, April 1980, 104-116

Variability in phonetic parameters (whether neuro-muscular, spatial or acoustic) is of two types. The first concerns how a same phonological segment (e.g. /p/) may be realised differently in different phonological contexts (e.g. /pæt/ vs. /tæp/); the second concerns how a same phonological segment may be realised differently on different occasions in the same phonological context (e.g. /pæt/$_1$ vs. /pæt/$_2$). It is this latter type to which we address ourselves in this paper.

Our overall model is simple and not original: a phonology (not of the competence-only type) produces an output object, X, according to two broad considerations:

the intention to encode as speech a sentence (itself the result of an enactment of the intention to encode a thought as language);

a set of static rules governing all such encodings (phonological competence).

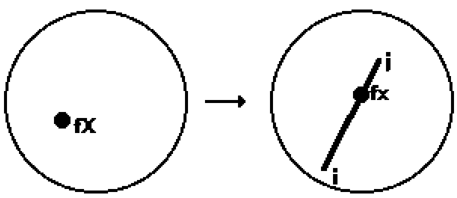
A second set of static rules governs how to employ b. to accomplish a. (And, we might add, other sets of static rules govern a-grammatical aspects of the encoding, such as phonological variants selected in accordance with social or stylistic, etc., constraints.)

X is an object which is psychologically real: that is, both a speaker and listener are aware of it as a linguistic entity having certain linguistically motivated attributes or features. Furthermore for the speaker and listener X is an invariant object. So, for example, /p/ in /pæt/$_1$ and /p/ in /pæt/$_2$ are the same object. A close examination of the phonetic output, x, deriving from X will reveal, however, that at the neuro-muscular, spatial and acoustic levels x varies in each of its constituent parameters. If each parameter is regarded as a unidimensional scale, then we have

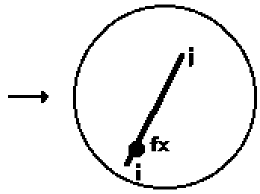
fX —» fx [*i* — *j*]

in which fX stands for a phonological feature of X which derives a phonetic feature fx of x, and in which *i*/*j* stands for the entire range of variability along the observed unidimensional scale of fx. We are interested in the characteristics (formal, not the mechanism *per se*) of —» and *i*/*j*. How (and why) is the realisation fx of fX constrained to fall within the range and not within some other (wider or narrower) range?

It is clear that fX is invariant. That is, fX is a *point* in some abstract psychological space. It is equally clear that any one fx is a point in some ‘real’ space, but that point falls on a defined line in that space — a line labelled fx[*i* — *j*]. Thus:



Whereas, for some other fx (i.e. fX realised on some other occasion), we have



Notice that •fx varies, but that [*i* — *j*] *does not*. And notice that •fx is constrained to vary only along [*i* — *j*] not to some other locus within the space.

fX —» fx [*i* — *j*]

defines [*i* — *j*].

fX —» fx

defines •fx.

Consider the following data which was obtained during a somewhat elaborate electromyography experiment (Daniloff, Morton and Tatham *forthcoming*). The phonetic parameter (fx) under consideration is contraction of *m. orbicularis oris* associated at the vocal-tract configuration level with (or deriving) bi-labial closure for /p/ in /pV.../ the vowel did not involve this parameter). Specifically the data scores refer to peak amplitude of emg signal obtained during the closure gesture and the units are arbitrary linear units of amplitude:

24	28	24
25	33.5	30
31	22.5	34
32	25.5	31
29	41	36
25.5	19.5	23.5
35	27	28
21.5	47	36
		29.5

We can see that the range of scores was 19.5–47, the arithmetic mean 30, the standard deviation 6.5, and the coefficient of variation 22.

Rarely stated, though implicit in almost all modern experimental literature, is the notion that somehow or other the arithmetic mean of such a set of data constitutes the aimed for target, equivalent to an unconstrained realisation, fx, of fX. This notion is, of course, a hypothesis which awaits empirical support or refutation. Whether such a claim is actually being made or whether, in fact, reduction of the data in this way is conducted simply to provide a single number to match up with a single symbol (like [p]) for the sake of convenience and thus duck the variance issue is not, however, usually clear: we prefer to believe that the latter is untrue and that many, many researchers are formulating some target hypothesis based on the arithmetic mean of a set of numerical observations.

Such a hypothesis is, of course, perfectly reasonable. It is rare to have very large data sets in phonetics (their collection may, for all sorts of reasons, be impossible), but it does seem that usually observed scores distribute fairly normally. And that fact supports the idea that the arithmetic mean describes a centre of gravity in the range of observations which might well constitute an ideal realisation of fX. We do not believe that this is a wholly satisfactory way of defining what have come to be known as targets: a better way, though quite out of the question experimentally at the moment, would be to examine the musculature's innervating signal at the moment it is generated in the motor cortex. For the time being, though, let us continue with the arithmetic mean calculated from observations made at the periphery. So, ideally (i.e. under conditions unconstrained by some shortcomings in the mechanisms involved), fX →fx (i.e. with no variation). The observed range is best, perhaps, accounted for as *departure* from fx, with explanation for that departure taking the form of constraints on —», to give fX →fx [*i* — *j*].

The coefficient of variation is of course just a single index taking in both the mean and the standard deviation to allow either comparison between variations about different means, or to predict variations, given a proposed mean. It predicts a range in which there is, say, a 90% chance for occurrence of any one fx. We have used the coefficient of variation as a measure of *precision* of articulation. This is based on the idea that a narrower range of variation indicates more precision in the realisation of fX.

Variation derives from the general inability of human beings to replicate tasks exactly (defined as: achieve at will the same score time and time again). What we notice, though, is that the coefficient of variation itself varies. On occasion it is wider or narrower than expected. Thus. lip-rounding for /u/ in English might, for peak contraction of *m. orbicularis oris* have a coefficient of variation 20, and /u/ in French 17, but /y/ in French 12. *Precision can be changed at will*. It follows that fX →fx [*i* — *j*] is too simple if → means 'directly derives'. An expression fX —*» fx [*i* — *j*] is more correct, and states that the derivation is a more-than-one-stage process.

Hence:

fX → fx [*p* — *q*] →fx [*i* — *j*].

The observed range, fx[*i* — *j*], derives from some naturally (i.e. intrinsic to the mechanism) determined range, fx[*p* — *q*]. The task is to explain the arrows, so let us distinguish between them:

fX —*u»* fx [*p* — *q*] —*l»* fx [*i* — *j*].

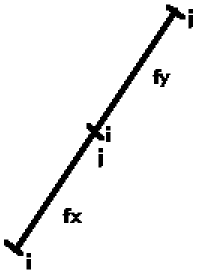
to have —*u»* and —l».

We have already said that —*u»* is some universally determined derivation: its explanation appears among the laws of myodynamics or mechanics, etc. —*l»* is not, however, universally determined; —*l»* is what causes the coefficient of variation of observed scores itself to vary. —*l»* cannot be universally determined if it can be changed at will. The explanation of —*l»* appears, we shall see, in the phonology.

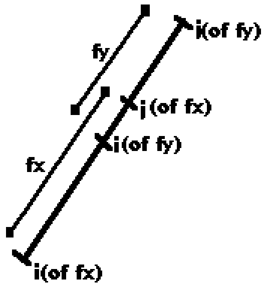
Under what conditions is the effect of —*l»* significant? We hypothesise that precision increases under certain 'social' constraints — when it seems necessary to 'talk more carefully'; we judge this to be linguistically trivial. But much more importantly we hypothesise that precision increases as *functional loading* on the parameter f increases. Thus, abstractly, if feature fl is one of several contributing to the unique identity of a psychologically real segment X (in the speaker or listener) — i.e. that X distinguishes from Y on more features than just fl — then the functional load (which could be given an index) on fl is relatively small. The limiting case occurs when X distinguishes from Y on *only* feature fl, in which case precision of phonetic realisation of the feature will be maximised.

In fact, though, staying with our example, functional loading on features varies itself, and is not simply stated by considering segments in isolation. Thus, in *j'ai dû* 'I had to' vs. *j'ai dit* 'I said' the functional loading approaches (or is) the maximum , since [j'ai dû] differs from [j'ai dit] on only one feature in one segment; *and* the two differently derived (semantically) phrases can easily co-occur. But [y] and it's lip-rounding feature are not so loaded in the phrase *du pain* 'some bread'. The sequence [dit pain] can only occur in *j'ai dit: 'pain'* 'I said: 'bread'', *vs. j'ai du pain* 'I've some bread'. But notice that it would be easy to say confusion is unlikely to occur here (*j'ai dit: 'pain'* being predictably rare). We have quite informally tried *j'ai d[i] pain* in a conversation discussing who has or has not any bread (!), and the occurrence of the 'wrong' segment was not noticed; i.e. a phonetically high-front *spread* vowel was decoded to a psychologically real high-front *round* vowel.

Functional loading at the phonological level is not however the only factor influencing precision of articulation. Phonetic features to not exhibit equal independence of controllability. It is generally taken that if a phonetic feature is independently controllable it can be used for realisation of some phonological feature. But in our earlier phonetic feature space it is not necessarily the case that fx[*i* — *j*] is completely distinct from, say, fy[*i* — *j*]. I.e. we may have:



It can easily be seen, then, that if *—l»* does not operate we might have



(since *—l»* has a restrictive effect on the outcome fx[*i* — *j*]. We hypothesise, then, that the restricting effect of *—l»* is there to prevent the potential overlap of the scales ** and ** , since such overlap would unduly jeopardise accurate decoding of fx[*i* — *j*] and fy[*i* — *j*] (the overlap region being ambiguous).

But the ambiguity itself is a *linguistic* phenomenon, and *—l»* (effectively a dis-ambiguator) is therefore a linguistic process based on knowing

how near to each other in the psychological space are fX and fY;

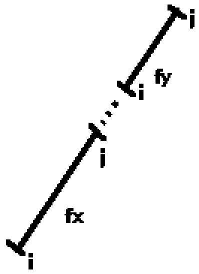
that fx and fy with their ranges of variation conjoin in a potentially ambiguity-producing fashion.

The linguistics would not, of course, have chosen to have fX and fY in the phonology's phonetic inventory unless it, in its turn, had known that *—l»* could disambiguate.

There are thus two aspects of independent controllability of phonetic features which make the candidates for phonological deployment (as derived psychological objects): direct controllability and governability of that control. If the governability of the control of fx[*i* — *j*] and fy[*i* — *j*] is insufficient to keep the unambiguous phonetically, then either fX may be selected or fY but *not both* (in any one language). Thus, we find $[i]_{Eng}$ and the (traditionally regarded as) slightly fronter, higher, tenser (phonetically) $[i]_{Fr}$, in English and French respectively. The model predicts that these could not co-occur in the same language as two distinct psychologically real objects.

When two such objects exist in different languages the psychological reality of the objects, we hypothesise, is often the same. Thus a French listener hearing $[i]_{Eng}$ in a French utterance will decode phonologically $/i/_{Fr}$. He may be aware that what he has decoded has had an 'English accent' (perhaps because the decoding process has been stretched, and he is aware of this), but phonological and other decoding will proceed, obviously, on the French /i/ and not the English /i/, although what triggered the $/i/_{Fr}$ may have been $[i]_{Eng}$. Since a speaker / listener of French cannot have two /i/s and neither can a speaker / listener of English, we say that there is only one /i/ (although there are two [i]s).

How narrow fx[i — *j*] and fy[i — *j*] can be made is critical. Suppose they can be mate this narrow:



then there is room for a fz[*i* — *j*] in between. And if f[*i*— *j*] can be made infinitely narrow then (psychological limitations permitting) there is space for an infinite number of phonetic targets. In fact, of course, we don't have to (formally) consider the psychological limitations here because the phonetic control limitations are grosser and define the number of f's possible. So, there is a range of possible *—l»* between

fX *—u»* fx[*p* — *q*] *—l»* fx[*i* — *j*],

where fx[*p* — *q*] = fx[*i* — *j*] (i.e. *—l»* does not govern), and

fX *—u»* fx[*p* — *q*] *—l»* fx[*i* — *j*],

where *i* and *j* are as close together as the control mechanism will allow (given that *i cannot* = *j*).

To summarise

The phonetic realisation of a psychologically real and invariant systematic phonetic segment is a two-stage process. Limitations inherent in the phonetic mechanism produce a variability in the final output which characterises the intrinsic precision of the system. Potentially a governing or stabilising mechanism may be employed to narrow that intrinsic precision. An occasion on which this occurs is when the phonetic output would promote ambiguity in decoding. The phonology's phonetic inventory consists of independently controllable phonetic features (raised to a psychological level) where this controllability subsumes intrinsic precision and governed precision. When the phonetic inventory has items requiring precision narrower than intrinsic, linguistically motivated governing is applied within the phonetics. Just as, to select normal independently controllable phonetic features, a phonology employs a model of intrinsic phonetic processes, so it must be aware of the governability of these processes.

Reference

Daniloff, R. G., Katherine Morton and M. Tatham (forthcoming). Monologue documenting some recent electromyography experiments.