Devoicing, Aspiration and Nasality — Cases of Universal Misunderstanding?

Katherine Morton
Mark Tatham

Reproduced from *Occasional Papers* 23, Department of Language and Linguistics, University of Essex UK, April 1980, 90–103.

Part of the task of linguistics is to throw light on what constrains human language to be the way it is. Some of the constraints are associated with the mind, some with the design of the vocal apparatus and its control. What is really interesting is when the mind manipulates constraints which are not absolute. [*Footnote:* It would be absurd to maintain that all constraints are absolute.] This paper tries to illustrate an area of constraint manipulation. Our task is relatively easy because we are working at the phonology/phonetics end of language — but we can imagine similar constraint manipulation at the semantics/syntax end.

Phonology is, for the most part, about the assigning of phonetic representations to surface structures derived in the syntactic component of a grammar (Chomsky and Halle 1968, p.7). This is taken as unarguable for the purposes of this paper. It is clear that a phonetic representation corresponds to some perceptual reality which itself may not relate in any direct way to the actual heard acoustic of an utterance. Accordingly the following rule would be quite wrong in a phonology of English:

1.* [C, +voice] —» [C, -voice] / V — V   * means ‘probably wrong’

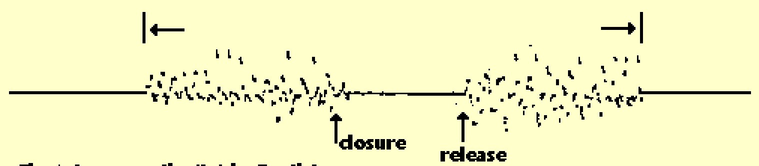


Fig.1 Intervocalic /b/ in English.
Note lack of vocal cord vibration during the stop phase.

There *is* a perceptual reality to the voice feature in intervocalic voiced consonants, as there is in most (all?) languages The fact that in our own dialects (respectively American and British English) many, if not most, speakers tend to fail to produce vocal cord vibration throughout the phonetic realisation of the consonant (Fig. 1) is irrelevant to phonology, and its irrelevance to perceptual phonetic representation is explained by the manner in which we perceive speech. Speakers of some languages (e.g. French) do however usually maintain the vocal cord vibration throughout the realisation of the consonant (Fig. 2) Inasmuch as the major relevance of the perceptual reality is its contrast with consonants having the perceptual reality of no voice, then that reality is in some sense identical for the speaker of English and the speaker of French. It is linguistically irrelevant that the acoustic cue triggering what is perceived differs between the two languages.

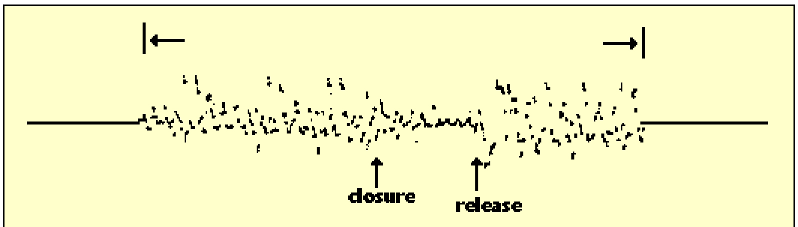


Fig. 2 Intervocalic /b/ in French.
Note continuing vocal cord vibration during the stop phase.

Similarly, but more contentiously:

2.* [C, +voice] —» [C, -voice] / — #

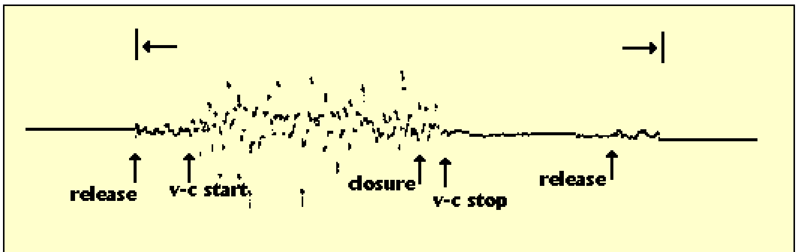


Fig. 3 English 'cab'

For all dialects of English vocal cord vibration is markedly absent during the greater part of the occlusion in the realisation of such consonants (Fig. 3), whereas in French the vocal cord vibration tends to continue until and often beyond the end of the occlusion (Fig. 4). Once again, though, for both languages there is a perceptual reality to the feature [voice] for such consonants. Notice that whereas Rule 1 is never found in descriptions of English phonology, Rule 2 quite often *is* found.

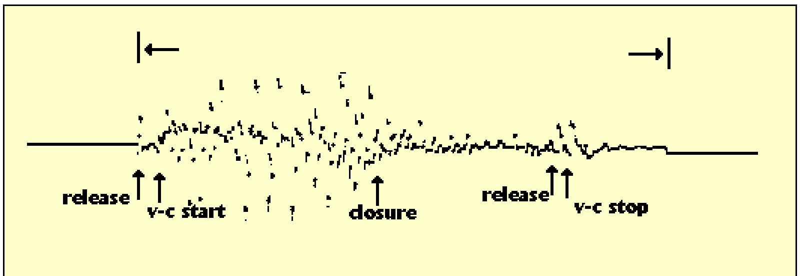


Fig. 4 French 'cab'

What might be the reason why phonologists have accepted Rule 2 but not Rule l? The prior existence of the rule:

3. [V] —» [V, +length] / — [C, +voice]

transferring in a sense the contrasting feature of the consonant to the preceding vowel (Fig. 5) might explain its (Rule 2's) inclusion in some phonologies of English: [+voice] on the consonant is no longer necessary, though it might be perceptually real. Is there some convention saying that if an otherwise contrasting feature has been made redundant (as, for example, by the inclusion of Rule 3), then, if not phonetically realised in a 'normal' (i.e. directly correlating) way, the feature should have its sign changed? If so, the psychological reality of the derived [+voice] feature on final underlying [+voice] consonants has not been captured, and the grammar fails.

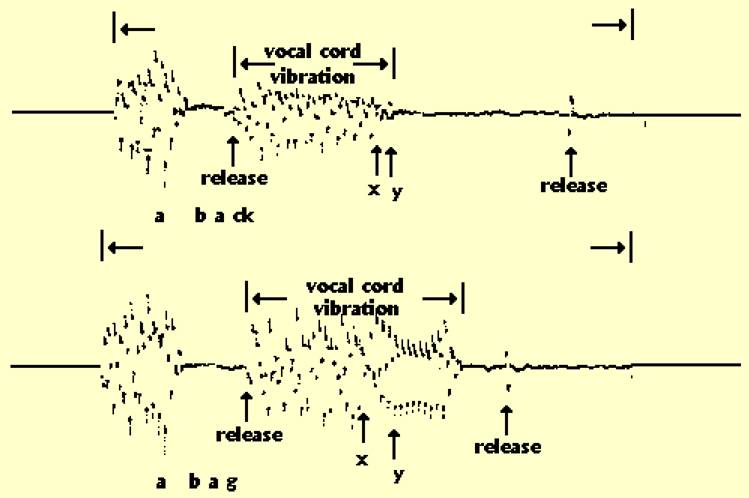


Fig. 5 'Vowel' lengthening before [+/-voice] stops.

Notice that we can say definitely that the period of vocal cord vibration has increased before the [+voice] stop. 5a is 'a back'; 5b is 'a bag' x is closure start; y is closure complete.

Similarly:

4.* [V] —» [V, +nasal] / [C, +nasal] — [C, +nasal] (Fig. 6)

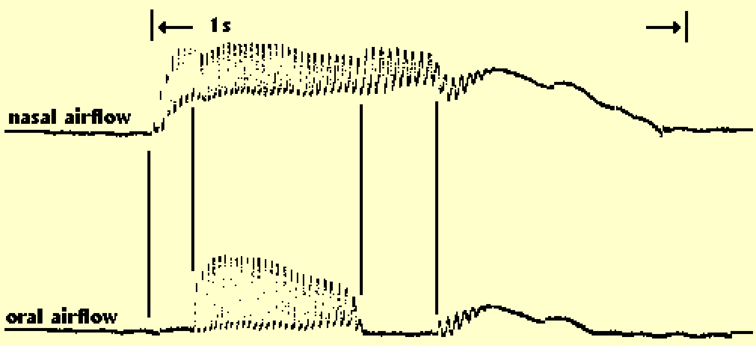
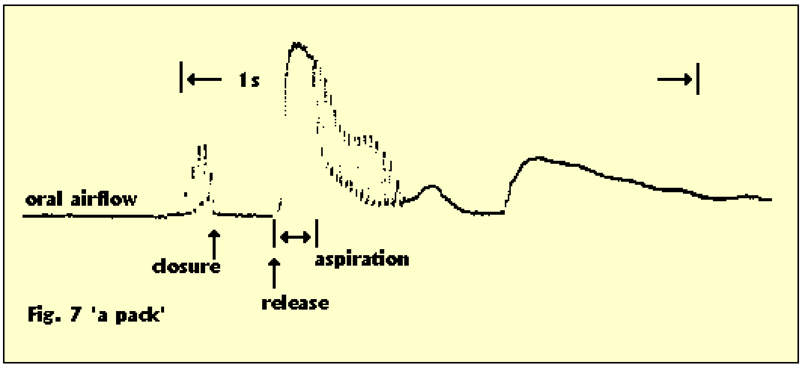


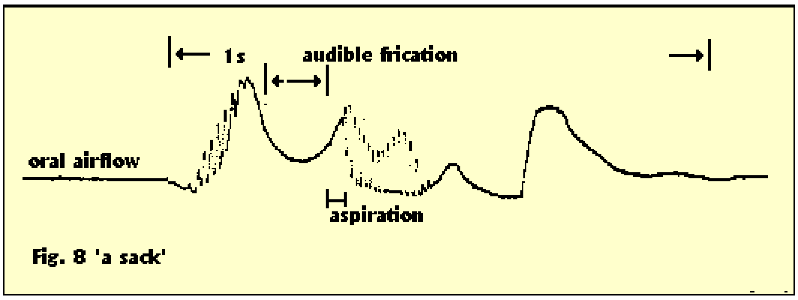
Fig. 6 British English 'man'

5.* [C, +stop, -voice] —» [C, +stop, -voice, +aspiration] / # — V (Fig. 7)



or, if some insist on having Rule 5, why not:

6.* [C, -voice] —» [C, -voice, +aspiration] / # — V (Fig. 8)



since the delay in onset of vocal cord vibration which is generally remarked following the realisation of [-voice] stops in English also occurs following [-voice] *fricatives* and affricates (though perhaps to a lesser extent for aerodynamic reasons)?

Take the other way of defining the notion of phonetic specification. The phonetic specification (the phonology's derived level) shall contain all linguistic information (required for subsequent listener decoding), and be such that some a-linguistic phonetics can deal automatically with its acoustic encoding. In other words, the phonetic specification amounts to an instruction set to enable a phonetics to proceed to realisation.

Cessation of vocal cord vibration in the realisation of a phonological [+voice] consonant in intervocalic environment is, of course, a non-programmed occurrence. By this we mean that it occurs as a non-intended artefact. Vocal cord tensing, raised subglottal air pressure, etc., are always present, and from this combination of phonetic features vocal cord vibration would normally occur spontaneously were it not for the fact that any oral occlusion will raise the supraglottal air pressure destroying the transglottal pressure drop which is conditional for vocal cord vibration. Vibration however can be, and is regularly by some speakers, continued for a little longer by reducing supraglottal air pressure to maintain the transglottal pressure drop. This is done by increasing the volume of the oral cavity, say by lowering the larynx.

Similarly in final position not only does the supraglottal air pressure rise because of the occlusion, but probably also the subglottal air pressure falls with the result that vocal cord vibration stops early in the consonant realisation. Once again, lowering the supraglottal pressure and at the same time maintaining (or raising) the subglottal pressure will prolong the vibration (as in done in, say, French).

Since more often than not in English vocal cord vibration falters in these two environments it is clear that a prolonged vibration cue is not required to establish perceived voicing on such segments. Notice that in both environments (intervocalic and final) the following *phonetic* rule holds:

7. (phonetic) [C, +voice] —» [C, <vocal cord vibration] / {V — V, — #}

In this type of phonetic rule the object to the left of the arrow is a phonological segment input to a phonetics, and the environment is also a phonological description. The symbol < has been used to mean ‘less than fully realised’.

Notice than an alternative rule (which would follow from Rules 1 and 2 above) is inappropriate for *underlying* [+voice] consonants:

8. (phonetic) [C, -voice] —» [C, -vocal cord vibration] / all environments

since it fails to characterise the facts. The output realisation of *no* vocal cord vibration obtains only when the vocal cords are insufficiently adducted, and occurs only when the input derives from a phonologically underlying [-voice] consonant (such as underlying /p/ or /s/, etc.). This rule cannot be right for inputs derived fro underlying [+voice] consonants. To prevent the operation of Rule 8 underlying [+voice] segments should not be described as [-voice] at the (phonological) derived level.

In some languages Rule 2, followed by a modified phonetic Rule 8 (restricting the environment to final position) would be correct. German and Russian are examples of this (Fig. 9). For those languages the [voice] distinction is neutralised at the phonological derived level and has no validity at any psychologically real phonetic level. In the phonetics of these languages vocal cord and larynx adjustment is probably identical for consonants marked [voice] and [-voice] at the underlying level of the phonology in final position.

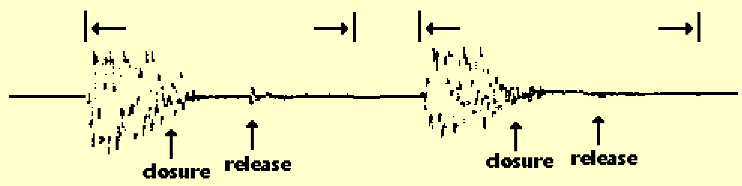


Fig. 9a German /bunt/ and /bunt/ (from: 'Bunt' and 'Bund')

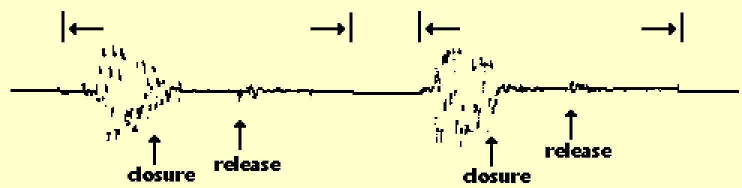


Fig. 9b Russian /Xljep/ and /Xljep/ (from: /Xljep/ and /Xljeb/)

But, if Rules 7 and 8 are correct, what then of French? If at the derived level underlying [+voice] consonants are [voice], then Rule 7 will derive an incorrect phonetic output, since French has (generally) full vocal cord vibration in these environments. It therefore follows that for French either Rule 7 must be prevented or what is to the left of the arrow must not be simply [C, +voice]. The latter solution is unacceptable since the psychological reality of the phonetic description is probably appropriately described as [C, +voice]. Rule 7 is however a universal phonetic rule and, unlike phonological rules, such phonetic rules cannot be simply omitted for particular languages: the rule describes an automatic coarticulatory effect explained by aerodynamic and mechanical laws which are independent of human behaviour. The instruction ‘Do not operate Rule 7’ is not a possible instruction. The only possibility is to inhibit the rule’s operation by changing the aerodynamic or mechanical conditions. Hence the observed voluntary lowering of the larynx to counteract air-pressure build-up, temporarily minimising Rule 7’s effect. French does this; English does not. For both the psychologically real phonetic segment is [C, +voice]. French, then, has the Production Instruction ‘Inhibit Rule 7’.

Such an inhibit instruction is not an additional rule, since

- it is not a phonological rule (i.e. not a language-specific voluntary rule deriving something perceptually and psychologically real);
- it is not a phonetic rule (i.e. does not characterise a universal phonetic process);
- it is not involved in linguistic decoding (i.e. does not derive some psychologically real segment from the acoustic — that would happen, as in English, without its operation).

Similarly, for both British and American English the perceptually real specification of a vowel between nasal consonants is [-nasal]. There is however a phonetic rule:

9. (phonetic) [V, -nasal] —» [V, +nasal] / [C, +nasal] — [C, +nasal]

That is, a derived level [-nasal] vowel occurring between two [nasal] consonants will when realised become phonetically nasalised. This rule describes a universal effect manifested as a partial (or delayed) stopping off of the nasal passage by the velum in such environments. America English however, in some dialects, see s to have nasalisation in excess of the normal coarticulatory effect. This may be the opposite situation to the voicing in French. Hence, a Production Instruction ‘Enhance Rule 9’.

These Production Instructions are language specific and serve to adjust the effect of universal phonetic rules where such adjustment is possible. They are neither fully linguistic since they to not assist in deriving the psychologically real phonetic specification — i.e. not lie, for example, so o rule deriving palatal and velar /l/ in English), nor are they fully phonetic (since they are language specific and do not characterise some universal phonetic process).

Finally similarly, the delayed onset of voice following an initial voiceless plosive (‘aspiration’) in English and other languages is described by the phonetic rule:

10. (phonetic) [V, +voice] —» [V, <vocal cord vibration] / # [C, -voice] —

where [<vocal cord vibration] means not fully vibrating vocal cords or delayed vocal cord vibration. This rule is also explained by aerodynamic and mechanical effects having nothing to do with language.

English shares with French, Spanish, Italian, etc., the psychologically real specification for the vowel as [voice]. In these other languages, however, a language-specific Production Instruction inhibits Rule 10, resulting in an acoustic similar to that produced for English vowels following phonologically [voice] consonants.

This latter example is interesting because it can be observed that the inhibiting Production Instruction is applied to different effect in the three languages cited. This difference of degree of application is voluntary, consistent, and need not necessarily be applied. Thus French does have some vocal cord vibration onset delay, Spanish has consistently less, and Italian consistently yet less or no delay (Fig. 10). We repeat, though, that for 11 four languages the phonetic specification (derived level in the phonology) is of a vowel which is [voice].

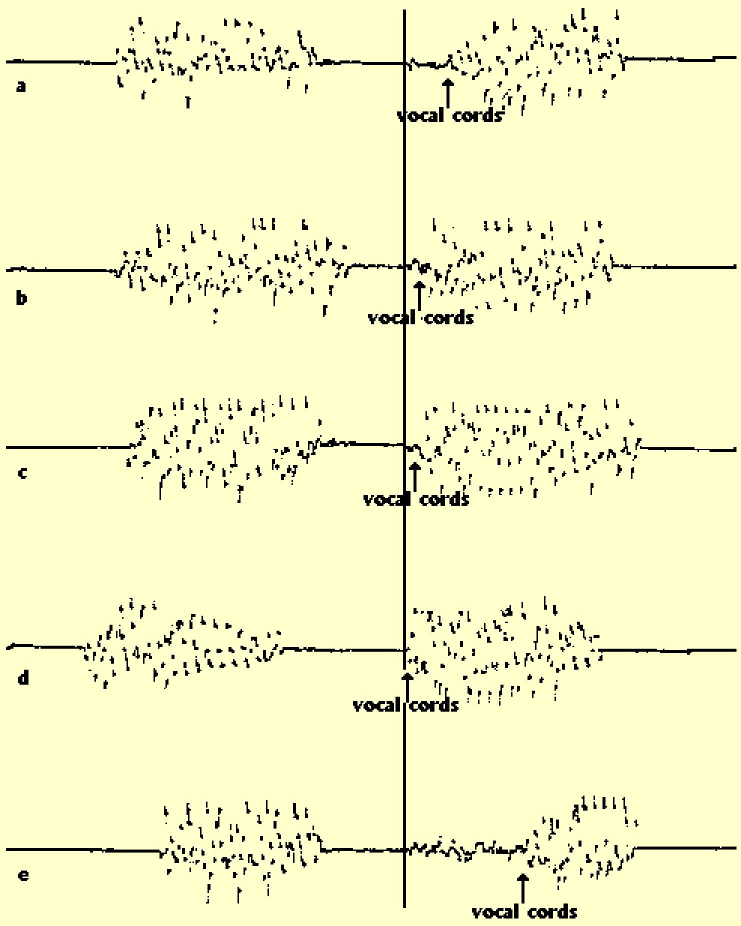


Fig. 10 Typical phonetic realisation of [-voice] stops (/p/) in:

a. English; b. French; c. Spanish; d. Italian.
Compare these with e. realisation in Korean of a [-voice, +asp] stop, where the phonetic feature has been raised to a contrastive phonological level.

But what, it might be asked, of a language like Korean, where vocal cord vibration onset timing in vowels following some consonants (always only stops) is phonologically significant This situation is analogous to the English vs. Russian example in which English derives surface variant /l/'s (palatal and velar) from a single underlying /l/, whereas Russian has the palatal and velar /l/'s at the underlying level. If an acoustic or articulatory phenomenon can be systematically controlled then there is no reason why it cannot be used at any level if it can be consistently perceived distinctively (i.e. can be given psychological reality). Delayed vocal cord vibration onset is used contrastively with non-delayed vocal cord vibration onset in languages such as Korean (whereas it is not in English — though see Hewlett's (1980) example of support vs. sport in fast speech), this in turn deriving from some underlying contrast. Similarly the version with full vocal cord vibration derives directly from some underlying representation. All things being equal this last type involving vocal cord vibration throughout the stop's closure phase is perhaps most likely to fail phonetically; yet because of the three-way contrast its correct phonetic realisation is critical.

The reason for saying that the fully-voiced stop is the most likely to fail is that it involves a production inhibition instruction raised to phonological status (and appropriately couched in abstract terms) to derive the full vocal cord vibration. The inhibition of the phonetic rule which would normally stop the vibration is however usually of limited application. We notice by casual observation only that the stops in Korean which maintain their vocal cord vibration invariably throughout the closure phase (in intervocalic environment) involve much less contact pressure for the closure than the other stops in the triplet. We have furthermore observed some air leakage through the closure during these stops. We hypothesise

- that force of contact for i. ‘aspirated’ stops and ii. ‘non-aspirated’ (but not voiced) stops will be more than enough to support the air-pressure build-up behind the place of closure (with non-aspirated aspirated, perhaps); and
- that for the ‘voiced’ stops force of contact will be just sufficient to maintain air-pressure build-up behind the closure up to a critical level corresponding to a transglottal pressure drop approaching the level at which vocal cord vibration will cease.

If the intraoral pressure rises further it conveniently leaks through the closure enough to keep the vocal cords vibrating, but not enough to destroy the ‘stop’ character of the consonant. We are currently doing the experiments. If we are right then we have an exciting example of very subtle phonological management of intrinsic phonetic phenomena. (Fig. 11)

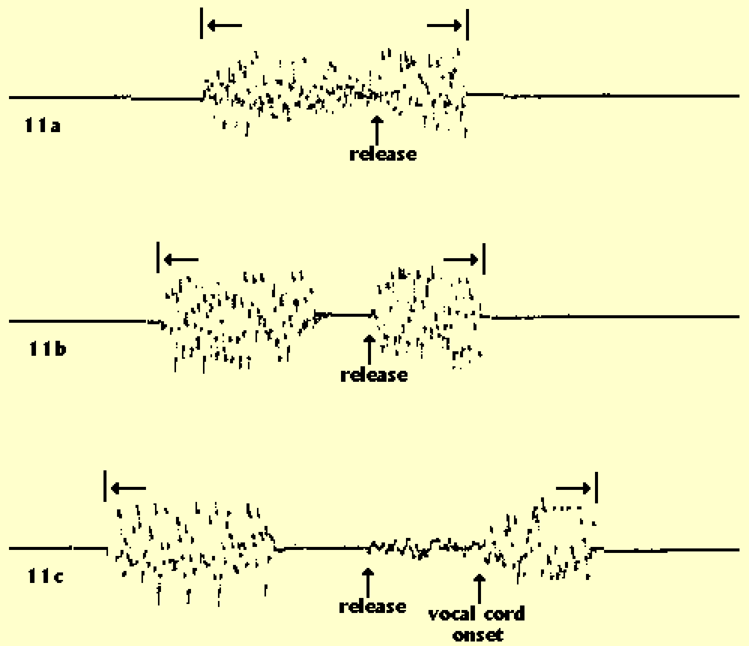


Fig.11 Intervocalic stops in Korean.
11a: [+voice, -asp]; 11b: [-voice, -asp]; 11c: [-voice, +asp].
The duration of each utterance is 500ms.

CONCLUSION

There exist phonetic coarticulatory phenomena deriving from properties inherent in the design of the vocal apparatus, its control, and the mechanical and aerodynamic processes involved. Many of these phenomena can be voluntarily influenced along a scale ranging from maximum possible inhibition (which may fall short of negation) through to maximum possible enhancement. Inhibition and enhancement are language-specific, though not necessarily phonological, since their operation may have no psychological reality for the speaker or listener. On occasion, and where consistent realisation is guaranteed, inhibition and enhancement can be raised to full phonological status, be language-specific and have psychological and perceptual reality. Our examples classify as follows:

Unmodified (?) coarticulation:

- lack of vocal cord vibration in phonologically [+voice] final consonants in English;
- failure of vocal cord vibration in intervocalic [+voice] consonants in English;
- nasal air-flow in British English during inter-nasal vowels;
- delay of vocal cord vibration in vowels following initial -voice] stops in English.

Inhibited coarticulation (non-phonological):

- short voice-onset-time associated with initial [-voice] stops in French, Spanish, Italian (differentially inhibited);
- vocal cord vibration during closure in inter-vocalic [+voice] stops in French.

Enhanced coarticulation (non-phonological):

- nasal air flow in American English during inter-nasal vowels.
- Inhibited coarticulation (employed phonologically):
- vocal cord vibration in intervocalic [voice, -aspiration] stops in Korean.

Enhanced coarticulation (employed phonologically):

- [?] over-delayed vocal cord-vibration in vowels following [-voice, +aspiration] stops in Korean (to maintain sure contrast with [-voice, -aspiration] stops.

Notice that the Korean example uses the entire inhibit — allow (or inhibit less) — enhance scale. It would be interesting to discover how many (for any given parameter) points on this scale could be used systematically for phonological purposes, maintaining perceptual distinction and therefore reality.

[As we go to press we have just received a copy of *Voicing in English and Finnish Stops* by Kari Suomi (*Publications* of the Department of Finnish and General Linguistics of the University of Turku; Turku: 1980) which gives an important and excellent summary of the literature concerning vocal cord vibration in [voice] stops. The model presented seems to us, however, to reiterate misconceptions about the relationship between phonology and phonetics. The systematic phonetic (output) level of the phonology is the final characterisation of what is psychologically real for the linguistically-naive speaker. What this level must interface with is the *input* level of the phonetics, not its output or some intermediary stage. Thus it is irrelevant to phonology that the phonetic output of realisations of /b, d, g/ may have variations in presence of vocal cord vibration: what is important to phonology is whether the conditions are set at the *initial* stages of the phonetics for vocal cord vibration — irrespective of whether (or how much of) this vibration occurs. It is definitely not the case that the observed output stands in a one-to-one relationship with the input; hence it is a mistake to take the output as a direct indicator of what is phonological. The output *less* the intrinsic effects, *less* the effects of the mediation of our Production Instructions is what relates immediately to what is psychologically real — and *that is invariant*.]