ED 132 824 FL 008 196

TITLE	Phonetics and Phonology. Occasional Papers, No. 16.
INSTITUTION	Essex Univ., Colchester (England). Language Centre.
PUB DATE	Apr 76
NOTE	97p.
AVAILABLE FROM	Language Centre, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, England. (limited supply, 1 pound, 4 pence)
EDRS PRICE	MF-$0.83 HC-$4.67 Plus Postage.
DESCRIPTORS	*Articulation (Speech); Consonants; Distinctive Features; French; *Generative Phonology; Language Research; Language Universals; Linguistic Theory; Morphophonemics; *Neurolinguistics; Oral Communication; Phonetic Analysis; *Phonetics; *Phonology; Pronunciation; Syllables; Vowels
IDENTIFIERS	Catalan

ABSTRACT

This volume is devoted to phonetics and phonology. It consists of the following papers: (1) "Generative Phonology, Dependency Phonology and Southern French," by J. Durand, which discusses aspects of a regional pronunciation of French, the status of syllables in generative phonology, and concepts of dependency phonology; (2) "On the Role of Notation and the Ordering of Rules in Phonology," by I.M. Roca, which discusses a system of notation in Catalan generative phonology; (3) "EMG - Techniques and Application to Speech Research," by C. Riordan, which describes a centralist model of speech production; (4) "Some Observations on the Role of Place and Manner of Articulation in the Perception of the Voicing Characteristics of Final Consonants," by D. O'Kane; and (5) "Orosensory Feedback Mechanisms and Speech Production" also by D. O'Kane. (AM)

University of Essex

Language Centre

OCCASIONAL PAPERS

Colchester, April 1976

16
PHONETICS AND PHONOLOGY

Occasional Papers 16

April 1976

Contents

GENERATIVE PHONOLOGY, DEPENDENCY PHONOLOGY AND SOUTHERN FRENCH

The purpose of this article[1] is threefold. First of all, it concentrates on some aspects of a major accent of French which has received little attention in recent years. Secondly, it adduces further evidence in favour of giving syllables a formal status withing generative phonology. Thirdly, it includes an illustration and defence of some of the concepts of dependency phonology as presented by Anderson and Jones (1974). In many respects the present acticle is an introduction to the phenomena in question and it is hoped that most of the ideas will be developed further elsewhere.

The accent which is described here is essentially the variety of French spoken in Hérault but the phenomena to be discussed seem to me to extend far beyond the limits of this "département" and perhaps to characterize Southern French as a whole[2]. The pronunciation

---

1. A preliminary version of this paper was presented in a departmental seminar at the University of Essex, and I wish to thank all my colleagues for their suggestions and support. A special debt of gratitude is owed to P.J.Brew, M.K. MacMahon and I.Roca, without whose constructive criticisms this article would not have existed. I am, of course, solely responsible for all remaining errors.
I have also made extensive use of my family and my friends in Hérault and elsewhere in the Midi and I wish to thank them for their cooperation and their patience.

2. My analysis corresponds for example to the description given for the Toulouse area in the classical study by Séguy (1951). Most of the Midi speakers I have encountered and studied seemed to conform to it. It would of course be premature to claim that this is the pronunciation of all Southern French speakers, nevertheless, the expression "Southern French" will be used throughout with a reference which, at this stage of our knowledge, should not be taken literally.

described here is a familiar colloquial one which partially explains why some of my assertions appear to clash with Martinet's and Deyhime's studies on the pronunciation of French. Deyhime (1967), to take the most recent approach, has made a careful questionnaire-survey of the pronunciation of a number of key-words for 500 subjects originating from various parts of France. His informants were first of all asked whether they made a particular distinction (e.g. between pot and peau, sotte and saute) or pronounced a word in such and such a way (e.g. did they say lion or buée in one or two syllables?). Deyhime also asked them to say the words in question and compared their actual pronunciation with the one they claimed to be making. Not surprisingly Deyhime's "Enquête sur la phonologie du français contemporain" produces what are in my opinion, rather odd results. 40% (34% to his ear) of his Midi informants would make a difference between sotte and saute, whereas I transcribe them both as [sɔtə]. Again 40% (34% to his ear) of the Midi informants declare that they distinguish jeune and jeûne, when I would claim that sociolinguistically such a person would acquire for his or her compeers a "prononciation parisienne" and would be felt not to belong to their linguistic community. But we know, and Labov (1966) has proved it beyond question, that there is an abyss between the way people speak and the way they think they speak. Furthermore, the pronunciation of people varies considerably when they read a list, a passage, or tell a story not to mention variations along other social parameters. I therefore feel that Deyhime's survey[3] to be useful, should be integrated to a sociolinguistic study such as the one made by Labov for New York English, and I will in what follows ignore his results.

3. Apart from the objections relating to the "reflexive" and artificial side of Deyhime's questionnaire which presupposes that theoretical terms like "syllable", "length", etc., have been clarified within linguistic theory let alone for naive informants, there are many problems connected with the social origin of his subjects (they are all students), the fact that they were all living in Paris at the time of the survey, and so on and so forth. Deyhime himself is aware of it (cf. his remarks pp 64-65) and it may be no accident that his final diagram (p.81) for Midi French matches my own observations (Part III).

The rules I will be concerned with are essentially low-level phonetic rules (in the sense of Schane (1968)), what would traditionally be called allophonic rules. The framework adopted here is that of generative phonology. In other words, it is assumed that there is no intermediary level between the (morpho-)phonological representations and the phonetic output. The reasons for rejecting such a level have been stated repeatedly in the specialised literature[4], and there would be little point in going over all of them once again. It may nevertheless be observed that most of the arguments given for French have been borrowed from morphological alternations rather than paradigmatic contrasts arising from identical underlying forms such as <u>rider</u> ~ <u>writer</u> in American English[5]. Not surprisingly fairly similar examples can be found within French. There are dialects of French, notably those comprised in the vast crescent that goes from the Ardennes to Normandy, where a contrast between /u/~/u:/, /y/~/y:/, /i/~/i:/, and /e/~/e:/ can be found[6]. The situation in this respect is the same as that which prevailed in XVIIth and XVIIIth century in France. Typical examples of this contrast are:

[1] la lie [li:] la boue [bu:] salée [sale:] nue [ny:]

le lit [li] le bout [bu] salé [sale] nu [ny]

---

4. See Chomsky (1964), Chomsky and Halle (1968). For a non-generative approach to phonology see Martinet (1965); a point by point rebuttal, as far as the treatment of French is concerned, can be found in Dell (1973a).

5. See Halle (1962).

6. Cf. Martinet (1969), pp 177-8 and 211-2, Deyhime, pp 70-1.

Given the tenets of autonomous phonemics, four extra phonemes should be recognized for these dialects since the contrasts in the same environment C ____ ≠ yield differences in meaning. But this approach seems to be neither economical nor insightful. It is obvious, at first sight, that the contrast occurs in a rather limited set of cases: feminine vs. masculine nouns, and feminine vs. masculine adjectives. We also know that feminine adjectives are normally formed in French by adding the neutral vowel /ə/ to the stem, this latter being dropped under certain conditions in most varieties of French[7]. It would therefore appear much simpler to posit that all the forms in [1] have an underlying schwa:

[2] la lie /liə/, la boue /buə/, salée /saleə/, nue /nyə/ and that a late rule[8] assimilates [ə] to the preceding vowel:

$$[3]\quad \begin{bmatrix} +\text{vocalic} \\ -\text{consonantal} \\ -\text{tense} \end{bmatrix} \rightarrow \begin{bmatrix} \alpha\ \text{back} \\ \beta\ \text{high} \\ \gamma\ \text{low} \\ \delta\ \text{round} \end{bmatrix} \Big/ \begin{bmatrix} +\ \text{vocalic} \\ -\ \text{consonantal} \\ \alpha\ \text{back} \\ \beta\ \text{high} \\ \gamma\ \text{low} \\ \delta\ \text{round} \end{bmatrix} \_\_\_\_ \neq$$

Rule [3] would convert the words of [2] into [4]:

[4] la lie [lii], la boue [buu], salée [salee], nue [nyy] which accounts for the contrasts we started from. Since schwa has in any case to be postulated for the proper functioning of French, all we need to add to the grammar is rule [3] which is a good example of regressive assimilation.

---

7. For the postulation of schwa in French consult Schane (1968), Dell (1973a), (1973b).

8. Schwa is at the phonetic level (once again in the sense of Schane (1968)) the only lax vowel. The features are those set out in Chomsky and Halle (1968), taking into account the suggestions made pp. 353-5.

It is on the basis of such arguments that the idea of a single autonomous level where all the contrasts could be established has been rejected. And as we indicated before we shall take as axiomatic that such a rejection is correct[9]. We shall, however, in what follows start from the "phonemic" system of standard French and compare it with the "phonemic" system of Southern French. The reasons for this are purely expository. All we need to assume is that the phonological segments mentioned below would appear at some point or other of the derivation and that they would undergo the rules we will presently mention (for further details on this see the Appendix).

The bulk of my argument (part III) consists in showing that the concept of the syllable[10] is called forth by our examples and allows more insightful generalizations. Part IV represents an attempt to solve a residual problem by having recourse to some formal notions provided by dependency phonology. Part II serves merely the function of refreshing the reader's mind for contrastive purposes.

II The Standard French System

At the "phonemic" level standard French possesses the following segments:

[5]

	TENSE SET				LAX SET
	-Back		+Back		
	+Round	-Round	+Round	-Round	
+high -low	y	i	u		ə = [-tense -back -high +low]
-high low	ø	e	o		
-high +low	œ	ɛ	ɔ	ɑ	

9. Notice that Schane (1971) proposes to recognize the phoneme as "a viable phonological unit" but says about autonomous phonemics that "The theoretical arguments against such a level are sound". (p. 503).

10. Notice incidentally that Schane (1968), in contradistinction to Chomsky and Halle (1968) uses the concept of the syllable for some of his phonetic adjustments. See, for instance, pp.35 and 42.

For some speakers an opposition is made between /a/ and /ɑ/ but it is losing ground and we follow Schane ((1968) p.19) and Martinet ((1969) p.188) in not seeing it as crucial in standard French. At the level we are envisaging /a/ is classified as (+back, -round), but would probably appear as central in the final phonetic representation. This does not pose any problem if we bear in mind the distinction between the classificatory and the phonetic function of features. It will also be noticed that we have not listed the nasal vowels as they are not central to our argument.

What should perhaps be mentioned is the fact that some of the oppositions high-lighted in (5) are weak or, to use Martinet's term, neutralized in many positions To quote Martinet (1969) at some length: "Entre Parisiens d'une même classe sociale, il n'y a aucun accord sur le degré d'ouverture de la voyelle de gai, quai, irai, les, ces, mes, etc. Ailleurs qu'a la finale, le choix de [ɛ] ou de [e] est presque automatique. Comme l'opposition [ø] - [œ] n'a jamais été d'une grande utilité, seul [o] - [ɔ], dont on voit mal comment on pourrait se dispenser et qu'une majorité de Méridionaux arrive à conserver, maintient la nécessité de distinguer, entre celui de /i/ et celui de /a/, deux degrés moyens d'ouverture du maxillaire" (p.189). To supplement Martinet's observations it can also be pointed out that the /o/ - /ɔ/ opposition is rather precarious in many contexts. It is sufficient to take the list of words ending in os and compare their transcriptions in a variety of dictionaries to persuade oneself that the situation is not absolutely straightforward:

[6]	JUILLAND	HARRAP	WARNANT
thermos	/os/	/ɔs/	/os/ & /ɔs/

As for Southern French making a general distinction between /o/ and /ɔ/ there is little point in resuscitating the objections made to Deyhime in the earlier part of this article.

III The Southern French System[11]

The Southern French System I am considering makes no "phonemic" difference between /ø/ - /œ/, /e/ - /ɛ/, and /o/ - /ɔ/. In other words, the following examples, which can be minimal pairs for standard French are pronounced in the same way:

[7] jeune [ʒœnə] été [ete] côte [kɔtə]

jeûne [ʒœnə] étais [ete] cotte [kɔtə]

Final schwa is normally pronounced before a pause (eg il est jeune ≠ ≠ [il e ʒœn] but dropped before a vowel (eg il est jeune et bête ≠ ≠ [il e ʒœn e bɛtə]. Let us characterize the system as follows:

[8] TENSE SET

	-Back +Round	-Back -Round	+Back +Round	+Back -Round
+high -low	y	i	u	
-high -low	ø	e	o	
-high +low				a

LAX SET

/ə/ = [-tense, -back, -high, +low]

11. Note that speakers of this type of French have an Occitan substratum and are in many cases in a situation of diglossia. See Lafont (1971) and Giordan (1975). The problem of relating this substratum (itself a complex set of interrelated dialects) to Southern French pronunciation is a fascinating one and useful to verify some theoretical assumptions but its magnitude makes it lie beyond the limited bounds of our study.

It is however the case that [ø] and [œ], [e] and [ɛ], [o] and [ɔ], all appear at the phonetic level. Here are some representative examples:

[9] (a)

heureux	[ørø]	mettez	[mete]	moto	[moto]
neutraliser	[nøtralize]	étriper	[etripe]	oprimer	[oprime]
Euclide	[øklide]	érable	[erable]	proclamer	[proklame]
creusiez	[krøzje]	étiolé	[etjole]	rosier	[rozje]

(b)

bonheur	[bonœr]	mer	[mɛr]	corporel	[kɔrporɛl]
secteur	[sɛktœr]	rester	[rɛste]	poste	[pɔste]
heurter	[œrte]	Elmire	[ɛlmire]	opticien	[ɔptisjɛ̃n]
oeil	[œj]	pareil	[parɛj]	roc	[rɔk]

(c)

neutre	[nœtre]	mettre	[mɛtre]	socle	[sɔkle]
pleure	[plœre]	guerre	[gɛre]	rose	[rɔze]

If for the moment we limit our attention to 9 (a) and (b) we can see that the high variants occur (i) before a word boundary, (ii) before a single consonant followed by a vowel, (iii) before two consonants provided that the second one is a liquid or a glide. On the other hand, the low variants occur (i) before a non-syllabic followed by a word boundary, (ii) before two consonants provided that the second one is not a liquid or a glide. The situation is summarised in [10].

[10] <u>High variant</u> [ø,e,o] <u>Low variant</u> [œ,ɛ,ɔ]

$$\text{High: } \_\_\_ \ \neq \qquad \_\_\_ \ C_1 \left(\begin{Bmatrix} L \\ G \end{Bmatrix}\right) V$$

$$\text{Low: } \_\_\_ \begin{Bmatrix} C \\ L \\ G \end{Bmatrix} \neq \qquad \_\_\_ \ C_1 \ C_2$$

where C_2 must be either an obstruent or a nasal, ie [+cons, -voc]

A superficial inspection of [10] shows that in each case the environments are completely disparate. What is it, for example, that makes a single consonant followed by a vowel have the same effect as a pause boundary? Why is [tr] different from [rt] (cf. neutraliser vs, heurter in [9][12]? It seems also purely accidental that two consonants of a certain type (eg [kt] in secteur) behave in the same way as a pause boundary. We can, of course, starting from our characterization of the system in [8] formulate a rule:

$$[11]\quad \begin{bmatrix}+\text{voc}\\-\text{cons}\\-\text{high}\\-\text{low}\end{bmatrix} \longrightarrow [+\text{low}] \;/\; \_\_\_\_ \;[-\text{syll}] \begin{Bmatrix} \# \\ \begin{bmatrix}-\text{voc}\\+\text{cons}\end{bmatrix}\end{Bmatrix}$$

but, as can be seen the choice comprised within the braces does not seem to be highly motivated[13]. Another solution might be to accept that syllables are part and parcel of phonological description. If we assume for the moment that the traditional

12. Saussure in his Cours (p.78) formulated very similar questions: "en vieux haut allemand hagl, balg, wagn, lang, donr, dorn sont devenues plus tard hagal, balg, wagan, lang. donnar, dorn ... Mais comment formuler la loi? D'où provient la différence? Sans doute des groupes de consonnes (gl, lg, gn, etc.) contenus dans ces mots. Il est bien clair qu'ils se composent d'une occlusive qui dans un cas est précédée, et dans l'autre suivie d'une liquide ou d'une nasale; mais qu'en résulte-t-il? Aussi longtemps que g et n sont supposés quantités homogènes on ne comprend pas pourquoi le contact g - n produirait d'autres effets que n - g."

13. Another possibility might be to take into account Lass's suggestion to assign the features [+cons, -voc] to word boundaries. This is a line of argument I will not pursue here.

way of assigning syllable boundaries in French is correct, a much clearer picture emerges. Let us consider the [ø] - [œ] examples of [9] (a) and (b), and rewrite them with syllable boundaries ($ indicates the beginning and/or end of a syllable):

12 (a)	$ ø $ r ø $	(b) $ b o $ n œ r $
	$ nø $ tra $ ll$ zə $	$ sɛk $ tœr $
	$ krø $ zjə $	$ œr $ tə $

If the same thing is done with the other examples of [9] (a) and (b) it will be obvious that the rule we are dealing with is a very simple and natural one. We get a high variant in open syllables (ie when no consonant intervenes between the vowel and the syllable boundary) and a low variant in closed syllables (ie when one or possibly more consonants intervene between the vowel and the syllable boundary). Instead of rule [11] we would now have:

[13]

$$\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ -\text{high} \\ -\text{low} \end{bmatrix} \longrightarrow [+\text{low}] \;/\; \underline{\qquad} \; [-\text{syll}] \; \$$$

and [13] is undoubtedly preferable to [11]. There is nevertheless a whole set of examples which still awaits discussion, namely [9] (c). All these words have a low variant in front of one consonant (or two consonants of a certain type) followed in turn by the neutral vowel. What is puzzling is that the two consonants in question are those that normally determine an open syllable (ie an obstruent followed by a liquid). And, in fact, in styles of standard French where schwa is pronounced (eg poetry) the accepted division of words like <u>neutre</u> and <u>guerre</u> into syllables is neu$tre and gue$rre. Why do we not obtain: [nøtrə] and [gerə]?

These examples may appear to be counter-evidence to a syllable-based phonological model. But notice that since the lowering of / ø,e,o / in [9] (c) occurs in environments different from those stipulated in [11] the rule is no simpler for non-syllabic phonology [14]:

[14]

$$\begin{bmatrix}+\text{voc}\\-\text{cons}\\-\text{high}\\-\text{low}\end{bmatrix} \longrightarrow [+\text{low}] \;/\; \_\_\_ [-\text{syll}] \left\{ \begin{matrix} \left(\begin{bmatrix}+\text{voc}\\+\text{cons}\end{bmatrix}\right) \begin{bmatrix}+\text{voc}\\-\text{cons}\\-\text{tense}\end{bmatrix} & \text{(a)}\\ \# & \text{(b)}\\ \begin{bmatrix}-\text{voc}\\+\text{cons}\end{bmatrix} & \text{(c)}\end{matrix}\right\}$$

If we bear in mind the normal conventions for the expansion of such schemata as [14] , we can see that [14] is in fact an abbreviation for four rules (given in terms of examples):

[14]'	(a)	(i)	neutre	[nœtrə]
		(ii)	veuille	[vœjə]
	(b)		mer	[mɛr]
	(c)		sectaire	[sɛktɛrə]

The reader will have noticed that we can modify [13] in exactly the same way as we modified [11], namely [15] (We shall hereafter refer to [15] as the FINAL ADJUSTMENT RULE abbreviated as FIN ADJ):

[15]
FIN ADJ

$$\begin{bmatrix}+\text{voc}\\-\text{cons}\\-\text{high}\\-\text{low}\end{bmatrix} \rightarrow [+\text{low}] / \_\_\_ [-\text{syll}] \left\{ \begin{matrix} \left(\begin{bmatrix}+\text{voc}\\+\text{cons}\end{bmatrix}\right) \begin{bmatrix}+\text{voc}\\-\text{cons}\\-\text{tense}\end{bmatrix} & \text{(a)}\\ \# & \text{(b)}\end{matrix}\right\}$$

The schema [15] abbreviates three rules and therefore, at the simple level of the formulation, represents a gain over [14]. Furthermore, it still embodies the insight that for the majority of words considered the variant obtained is determined by its position within the syllable.

On the basis of the evidence presented here we would therefore argue that the introduction of syllables would afford greater insights into phonological processes. It could of course be objected that such a modification of phonological theory is hardly warranted in view of the superficial nature of our examples. They are after all equivalent to allophonic variants within a taxonomic model of phonology. This argument is not flawless. First of all, there is a fair amount of evidence from various languages in favour of giving syllables a formal theoretical status (in particular see Hooper, (1972); Anderson and Jones, (1974)). Secondly, if we consider the grammar as a way of connecting sound and meaning, then all aspects of the sound structure of a language must be dealt with, in the same way that all aspects of its semantic structure must come under investigation. And, whatever level the linguistic phenomena belong to, arguments based on "economy" and "insightfulness" must be able to operate.

There is nevertheless an outstanding problem in the formulation of FIN ADJ (R 15). It is still true that 15 (a) (i) and (ii) represent exceptions to the more normal pattern. But if we believe in the naturalness of phonological processes we may, once again, wonder what it is that makes one consonant (or two consonants of a certain type) followed by a schwa have the same effect as a closed syllable. Or to put it another way, what is it that overrules the syllabic bracketing in words like autre (au$tre), guerre (gue$rre) or pleure (pleu$re)? It cannot be a mere coincidence that the vowel which seems to bring about this realization, ie [ə], is unique at the low phonetic level[14]. It is the only member of the lax set, as [8] emphasizes, and therefore always unstressed. In part IV, I shall attempt to provide some

---

14. This is not true at the phonological level, since the latter is composed of five tense vowels and eight lax ones including schwa. See Schane (1971) p. 346 and the appendix.

explanation for this phenomenon[15]. In order to do this, I have had to approach this topic from the angle of dependency phonology.

IV SYLLABLES AND DEPENDENCY PHONOLOGY

One of the main problems that have always beset phonologists dealing with syllables has been that of assigning syllable boundaries and establishing criteria for so doing. Is it, for example, absolutely certain that a word like rabbit should be analysed as ræ$bɪt? Various phonologists have pointed out that syllables appeared to be "interlocking" units but unfortunately this has generally led them to the rejection of the syllable as a viable phonological entity. To put the discussion at a more technical level, this means that phonologists have been reluctant to allow non-proper bracketing of syllables. By a proper bracketing, is meant one that assigns a given element to one and only one constituent. At the level of syntax, given three formatives A, B and C, the surface structure A B C cannot specify A B as a phrase and B C as a phrase: the string may be bracketed as [[AB]C] or [A[BC]] but not both ways simultaneously [A[B]C]. To return to our previous example rabbit /ræbɪt/, the medial /b/ must belong to the first syllable or to the second one but not to both at the same time.

15. Observe that the lowering cannot be due to vowel harmony (or, to use another terminology, regressive assimilation), since /e/ is [-high, -low] in état [eta], /o/ [-high, -low] in otage [otaʒə], in spite of the following /a/ [-high, +low]. Nor can it be brought about, as could be argued in standard French (see eg, Schane (1968), p.47), by dropping of schwa, (which creates a closed syllable), since in that dialect schwa is generally not dropped. This argument is in any case dubious even for standard French since in the styles where the neutral vowel is retained the lowering still takes place.

Anderson and Jones (1974) have recently put forward a number of arguments re-affirming the relevance of the syllable as a structural unit, and showing that this was possible if one were prepared, at least for languages like English, to permit non-proper bracketing. The criterion they suggest for assigning boundaries to medial clusters is to take into account the constraints on final and initial clusters of monosyllables. Let us exemplify this with the help of two cases. A word like <u>rabbit</u> would be bracketed as follows: $[_1$ræ$[_2$b$]_1$ɪt$]_2$ (the square brackets designate syllable boundaries not a phonetic transcription) with the /b/ belonging to both the first and the second syllable. The reason of this is simple: /b/ can be syllable final in a monosyllable (cf. <u>cab</u>) and it can also be syllable-initial (cf. <u>bit</u>). To take a more complex example, <u>reluctant</u> would be bracketed as $[_1$rɪ$[_2$l$]_1$ʌk$[_3$t$]_2$ənt$]_3$. The interesting cluster is the second medial cluster <u>ct</u>. We know that /kt/ is possible as the end of a monosyllable (cf. <u>duct</u>) but we also know that there is no monosyllable beginning with /kt/, therefore whereas the second syllable encompasses /kt/, the third one can only begin with /t/, since /t/ is possible initially (cf. <u>tick</u>). Anderson and Jones (1974) demonstrate that such a treatment clarifies a number of phonetic and phonological facts (particularly relating to stress rules in English) which otherwise remain unexplained.

They then go on to claim that the following principles govern the character of medial clusters (p.7):

[16] (a) Medial clusters are composed of a syllable-final preceding a syllable-initial sequence.

(b) "Precede" includes overlap where possible.

Interestingly enough, Anderson, in a short but thought-provoking note ((Anderson) 1973), had broached the same question in a more universalistic spirit. He speaks there in terms of "interludes","codas" and onsets" and formulates the governing principles thus:

[16] (a) Interlude = coda ≦ onset

(b) = is preferred to <

("≦" indicates precedence, "<" strict precedence, "=" coincidence).

Anderson adds that such a rule "could make a fair claim to being a language universal, provided the coda is allowed to be null". He also declares that "French, not surprisingly, conforms to [16] " and adduces some evidence in favour of overlapping syllables in that language. Although this decision may be correct, it seems as yet difficult to see what advantages could be drawn from it. It would appear that the same generalizations can be achieved whether or not we permit overlap in French. Furthermore, the tradition of French phonologists is overwhelmingly in favour of assigning single medial consonants to the following vowel, not to mention the well-known phenomenon of final consonants jumping across morphene or word boundaries (cf. petit ami = pe - ti - ta - mi). And this partially accounts for the discrepancy that exists in French between the phonology and the morphology. To quote Garde[16]((1968) pp.23-4): "Ainsi la syllable n'est pas incluse dans le mot en français, mais elle l'est en russe, et dans beaucoup d'autres langues: la coupe de syllables coincide avec la coupe de mots en russe mais non en français: comparez "choeur angélique, coupe syllabique [k'œ-ʀɑ̃-ʒe-l'ik] et r, x'or 'angelov, coupe syllabique [x'or-'an-g,i-ləf][16]. I shall assume that this point of view is correct and assign medial consonants unambiguously to the preceding or following syllable. The only constraint that will be taken into account will be based on permissible intial sequences.

16. See also Matthews (1974), p.166.

Alternatively, since this decision raises some problems which we cannot go into here[17], we can for present purposes predict syllable boundaries by a rule such as the one presented in Hooper (1972) (but not envisaging it as a universal):

[17] $$\emptyset \longrightarrow \$ \ / \ [+\text{syll}] \left\{ \begin{array}{l} \underline{\quad}\ [-\text{syll}]_0^1 \\ [-\text{syll}]_0 \ \underline{\quad} \left\{ \begin{array}{l} [-\text{son}] \begin{bmatrix} +\text{son} \\ -\text{nas} \end{bmatrix}_0 \\ [+\text{cons}][-\text{cons}]_0 \end{array} \right\} \end{array} \right\} [+\text{syll}]$$

Either decision, perhaps with some modifications, will allow FIN ADJ [R15] to operate.

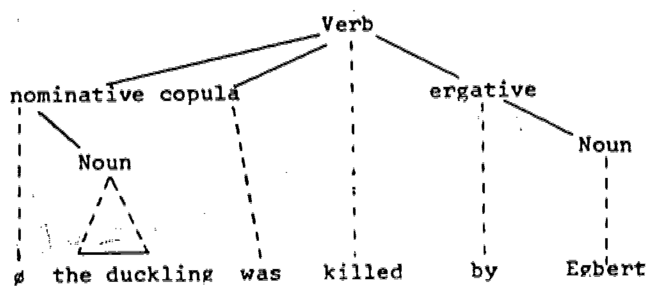
The next problem we need to deal with is that of the structural representation of syllables. Implicit in most traditional approaches is the idea that constituency rules of the form: Syllable $\longrightarrow$ (C) V (C) indicate quite adequately the structure of syllables. Note that this suggests among other things that consonants and vowels belong to the same rank. Once again Anderson and Jones make proposals of quite a different nature: they put forward the case that syllables could be more appropriately represented by dependency trees. Dependency trees have already been used in syntax by a variety of linguists (including Anderson)[18]. In a dependency tree pre-terminal categories have been eliminated, and in place of the constituency relationship (eg Det + Noun = Noun-Phrase), the categories

---

17. Particularly in relation to interludes the first element of which is /s/ (Cf. rester [rɛste]), and which according to our criterion should be re$ster (yielding *[reste]) since /s/ is possible initially in a cluster. Cf sport, stable, etc., etc.

18. On this problem see Anderson and Jones (1974) pp 9-15. See too Anderson (1971) Ch II, and the references therein. The rest of our paper borrows rather heavily from Anderson and Jones (1974), we can only advise the reader to turn to this article for further explanations or clarifications.

are "hierarchized" with respect to dependency. The dependency relationship itself is related to traditional notions like government, rection, rank and also endocentricity. Anderson (1971, p.43) within the framework of his localistic grammar of case, represents the sentence "the duckling was killed by Egbert" in the following way:



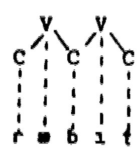
The position of Verb indicates its hyper-relational and governing function within the sentence; as for the case elements they express the relation contracted between the dependent Nouns and the governing Verb. And as is immediately obvious each category has a surface realization (ø could be interpreted as the deletion of the preposition of in subject position: cf. the killing of the duckling).

Anderson and Jones extend this type of representation to syllables. Each syllable has for central governing element a syllabic segment, and all the other segments are subordinate to it. If we may be forgiven for quoting Anderson and Jones (1974, p.9) rather at length: "such a characterization of syllables is appropriate if we can provide an empirical interpretation of subordination such that for each syllable there is a determinate centre ... to which all other elements are subordinate. We claim that syllabicity, and more generally, degree of sonority, are just such an interpretation, and that the syllabic element is the centre of each syllable. Associated with each syllable there is a unique sonority peak whose identification is available to native speakers, and enables them, for instance, to count the number of syllables in an utterance". If we take a syllable like bit /bɪt/ there is a syllabic /ɪ/ to which all the other elements are subordinate. The graph structure we would associate with it is as follows:

[18]

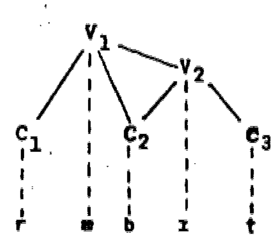
The word rabbit /ræbɪt/ which as we said was bracketed as [ræ[b]ɪt] will correspond to the tree [19]:

[19]



The /b/ is bidependent on /æ/ and /ɪ/ since, as our bracketing shows, it belongs to both the first and the second syllable. Anderson and Jones then go on to argue that one of the functions of the stress assignment rules is to select one of the syllabics of a word and to subordinate the others to it. The word rabbit would now be represented as follows:

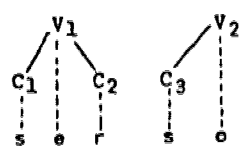
[19]'



At this point we need to introduce a theoretical distinction which is going to be crucial to the rest of our argument. If there is an arc leading from a particular node (say C_1) to another node which is higher in the tree (say V_1) without there being any intervening node between them we shall say that the lower node is dependent on the first one. In the case of [19]: C_1, C_2, V_2 depend on V_1; C_2 and C_3 depend on V_2 but C_3 does <u>not</u> depend on V_1. If on the other hand there is <u>a</u> path leading from a particular node to a higher node (irrespective of the presence of other nodes on this path) we shall speak of subordination. C_1, C_2, V_2, C_3 are therefore all subordinate to V_1. Having equipped ourselves with these theoretical tools we can now turn to our French problem.

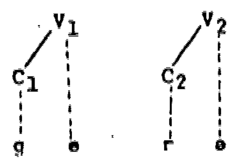
I have suggested earlier that it might be preferable, until further evidence is found, not to allow interlocking syllables in French. A medial consonant (whether part of a cluster or not) will belong to one and only one syllable. If we take for example the word <u>cerceau</u>, at the level of derivation in question, its phonological representation will be /serso/. It will then be bracketed $[_1 ser]_1[_2 so]_2$ (or $\$ser\$so\$$, if we wish to keep the same symbols). If we now choose to represent phonological words by means of dependency trees all French words will appear as a series of disconnected graphs. The word /serso/, to use the same example, will appear as [20]:

[20]



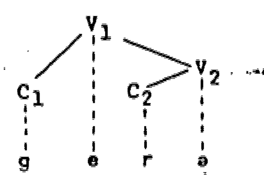
FIN ADJ [R15] will operate at this point and convert /serso/ into [sɛrso], since the bracketing of the word indicates that the first syllable is closed and therefore that the environmental conditions for 15 (b) are met. Until now it could be said that the use of dependency trees is nothing other than a notational variant of a traditional structural representation. This is also, at least at first sight, true of the contentious examples of [9] (c). Let us take the word guerre, its "phonemic" structure is /gerə/ and according to the criteria mentioned above for the division into syllables it must be bracketed as $[_1 \text{ge}]_1\ [_2 \text{rə}]_2$. The dependency tree corresponding to this is:

[21]



The situation seems in no way improved. Why should there be a lowering of the first vowel in this particular configuration. But notice that guerre is stressed on the first syllable, and that the second syllable is unstressed. In other words, we need, according to the conventions established, to subordinate the less or non-stressed syllabic to the stressed one:

[22]



A cursory glance at [22] reveals some connection between C_2 and V_1; there is a path leading from V_1 to C_2. The same is true of the other words of [9] (c) which are all stressed on the first syllable. As we said above, when a segment (eg C_1 in [20]) is directly dominated by the governing element we will describe it as a dependent element; if it is simply dominated (eg C_2 in relation to V_1 in [22]) we will speak of subordination. We could therefore propose, instead of FIN ADJ, the following tentative rule (handling as typical cases [20] and [22]):

[23] A [-high,-low] vowel segment becomes [+low]in the context of a right-hand side dependent or subordinate consonant.

This formulation would correspond to the intuitive feeling that in words like <u>pleure</u>, <u>guètres</u>, <u>autre</u> the consonant(s) somehow belong to the first vowel in spite of the syllabic structure. But it is also the case that, if we compare the right-hand side environments of the crucial vowels in [20] and [22], another generalization is possible: both V_1's have a dependent element (in one case a consonant, in the other a vowel) to the right. In consequence, another simpler rule is possible:

[24] A [-high,-low] vowel segment becomes [+low] in the context of a right-hand side dependent segment.

We have now found a rule that covers all the cases of [9], (b) and (c) without treating [9] (c) as a set of exceptions. This rule is undeniably more insightful and simpler than FIN ADJ, although it could be argued that the introduction of such complex machinery is perahps not justified by the generalization we were looking for. I would accept this criticism and hope that the examination of a wider range of data will help to clarify the issue.

Before I close the discussion, I would like to look back to the words of [9](c). The reader will have noticed that all of them have a schwa in final position. It is not impossible, however, to find examples where the neutral vowel is not final:

[25] aigrelet, osselets
formellement, autrement, seulement
rêverie, sonnerie, pleutrerie

In Southern French the neutral vowel is never dropped in that position. The examples of [25] would be phonetically realized as:

[26] [ɛgrələ], [ɔsələ], [fɔrmɛləmãŋ], [ɔtrəmãŋ], [sœləmãŋ], [rɛvəri], [sɔnəri], [plœtrəri].

The words of [25] are all composed of a stem + /ə/ + a derivational suffix.[19] According to Schane's convention (1968) pp 62-3, they should all be stressed on the suffix. But the stem of these words must also be stressed since the [œ] of <u>seulement</u> and the [ɛ] of <u>formellement</u> are issued respectively from /ɔ/ and /a/: fronting takes place under the influence of stress. We could therefore put forward the following informal rule:

[27] When the suffixes <u>-let</u>, <u>-ment</u>, <u>-rie</u>, <u>-ron</u> are added to a stem plus /ə/, both the suffix and the stem are stressed.

Rule [27] will enable our new formulations of FIN ADJ (ie rule [24]) to operate. But the fact that [ə] is pronounced in this position in Southern French has some unfortunate consequences. A word like <u>formellement</u> according to [27], must be stressed as follows:

[28] 0 1 0 1
formellement

But it is clearly undesirable that the first vowel in [28] should be as unstressed as schwa[20]

19. I have been able to find only one example that was clearly not derivational, namely céleri [səleri].

20. It may be objected that it is , in any case, misleading to speak of stress when all we are dealing with is different vowel qualities. Such an objection is based on the premise that stress is only associated with positive properties (eg loudness). There are reasons to doubt that such a view is correct: see Chomsky (1968), Ch III and, within a "functionalist" framework, Garde (1968), pp 57-66.

The difficulty we are now encountering is that of dealing with stress at the low-phonetic level. Schane has convincingly demonstrated that since many of the phonological changes are a consequence of stress, stress placement has to be a relatively high-ordered rule and must, of course, apply at the underlying level. There is no one-to-one correspondence between stress at the (morpho-) phonological level and stress at the phonetic level where the last syllable of a word (unless a schwa follows) must be accentuated. But it is arguable whether stress is a property of words as such at the phonetic level. It is indeed true that in citation forms the last syllable of a word is always stressed. But this rule is only a subset of all the rules that operate at the utterance level. Another rule is, for example, that in NP-VP constructions, the last syllable of each phrase is stressed. Natal'ie dorm'ait, la pet'ite dorm'ait, but la petite f'ille dorm'ait. To allow the operation of these rules some of which have a clear syntactic basis, we could suggest the following convention:

[29] At the low-phonetic level all tense vowels are stressed.

A word like formellement would, first of all, be stressed in accordance with Schane's rule and appear, let us say, in the form [28]. [29] which is a very late rule, would convert it to [30]:

[30]
```
1 1 0 1
formellement
```

And [30] would be the input to the stress-assignemnt rules at the utterance level. If it happens to be given in its citation form, the last syllable will be stressed (by convention all other stresses are reduced by 1):

[31]
```
2 2 0 1
formellement
```

On the other hand in [32], in a NP-VP construction formellement has the same stress contour as [30] except that all the stresses are reduced by 1.

[32] 2 2 2 2 1 2 2 2 0 2 2 2 1 0
La contestation est formellement interdite.

The advantage of [29] is that it sòlves the puzzling problem of the stress structure of [28], and that it allows rule [24] to operate should we discover more and more examples like celeri which do not conform to the morphological pattern of [25]. The reader may perahps think that at the end of the exercise things are not all that simplified, but as P. Anderson put it: "I have yet to see any problem, however complicated, which when you looked at it the right way, did not become still more complicated."[21]

---

21. Quoted in Katz (1972), p 56.

APPENDIX

In some respects our description may appear to simplify the actual situation in Southern French. If we assume that Southern French has the same underlying structure as Standard French, and there is some evidence that this assumption is correct, then some of the low-level phonetic realizations are directly specified by the morpho-phonological rules. For instance, to account for:

<u>clarté</u> - <u>clair</u>; <u>solitude</u> - <u>seul</u>

there would be a rule of the form:

$$\begin{bmatrix} +\text{low} \\ +\text{back} \end{bmatrix} \rightarrow [-\text{back}] \Big/ \begin{bmatrix} \overline{} \\ +\text{stress} \end{bmatrix}$$

In that case <u>clair</u> and <u>seul</u> would appear as [klɛr] and [sœl] without having to undergo our rule [15] FIN ADJ. Observe however that the first vocalic segment of <u>solitude</u>, specified as [+low, +back, +round, -tense] in the underlying representations would later have to be raised in the accent in question since the final representation is [solitude]. The raising would be triggered off by the fact that /ɔ/ is, in that case, in an open syllable. The same is true of a word like <u>bénir</u>, whose morpho-phonological representation is /bɛnir/ (in order to account for /bjɛ̃/ under stress), but the first vocalic segment of which would appear as [e] at the phonetic level. All these examples call for a rule of the form:

$$\begin{bmatrix} +\text{voc} \\ -\text{cons} \\ -\text{high} \\ +\text{low} \end{bmatrix} \longrightarrow [-\text{low}] \Big/ \underline{}\ \$$$

But notice that FIN ADJ in its present form is needed to account for the [e] and [o] alternations that we find in examples like <u>espérer</u> - <u>espère</u> and <u>sot</u> - <u>sottise</u> - <u>sotte</u>, all derived fr[om th]e underlying /E/ and /O/.

And if we take Schane's symbols seriously these must be [-high, -low, +tense]. (See Schane (1972), p.346). In the same way if we consider the alternation altitude - haut - haute, pronounced [altityde],[o],[ɔte] in our accent, the [ɔ] of haute will have to be derived from /o/. In other words, once again FIN ADJ will have to operate. Our case is therefore strengthened on the *prima facie* evidence of these examples, since we now need to specify the two environments of [10] (ie closed vs. open syllables). The introduction of syllables on grounds of naturalness and simplicity appears as even more desirable.

I have also left aside the problem of the formalisation of the dependency rules. A possible symbolic representation of rule [24] might be (where "$\lessdot$" indicates strict precedence and "$\Rrightarrow$" is the oriented dependency relation):

$$\text{SD:} \quad \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{syll} \\ -\text{high} \\ -\text{low} \end{bmatrix} < [+\text{segment}]; \quad \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{syll} \\ -\text{high} \\ -\text{low} \end{bmatrix} \Rrightarrow [+\text{segment}]$$

$$\text{SC:} \quad \begin{bmatrix} +\text{voc} \\ -\text{cons} \\ +\text{syll} \\ -\text{high} \\ -\text{low} \end{bmatrix} \longrightarrow [+\text{low}]$$

("Strict precedence" is not to be interpreted as "comes immediately before" but in opposition to "precedence" $\leqslant$.) (For further details see Anderson and Jones (1974)).

REFERENCES

Anderson, J.M:	(1971) The Grammar of Case. Towards a Localistic Theory, Cambridge University Press.
	(1973) A wee note on interludes and truncation in French. Edinburgh Working Papers in Linguistics. No.2.
Anderson, J.M., & Jones, C:	Three theses concerning phonological representations. JL Vol 10, No 1, pp 1-26.
Brew P:	(MS) Some pronunciation and spelling rules of French.
Chomsky N:	(1964) Current Issues in Linguistic Theory. In The Structure of Language. Fodor, J A & Katz, JJ (eds). Prentice Hall.
Chomsky N & Halle M:	(1968) The Sound Pattern of English. New York: Harper and Row.
Dubois, J. et al	(1971) Dictionnaire du français contemporain.
Dell, F:	(1973a) e muet, fiction graphique ou réalité linguistique. In a festschrift for morris halle. Anderson, S R and Kiparsky, P (eds). New York, Rinehart and Winston.
	(1973b) Les règles et les sons: introduction à la phonologie générative. Paris: Hermann.
Deyhime, G:	(1967) Enquête sur la phonologie du français contemporain. La Linguistique, III, Partie 1, pp 97-118, partie 2, pp 57-84.

Foley, J: (MS) Systematic Morphophonology

Garde, P: (1968) L'accent. Paris, PUF.

Halle, M: (1962) Phonology in Generative Grammar. Word, 18. pp. 54-72.

Word, 18. pp 54-72.

Hooper, J B: (1972) The Syllable in Phonological Theory. Language. Vol. 48. No. 3. pp. 525-40.

Juilland, A: (1965) Dictionnaire inverse de la langue française. The Hague: Mouton.

Katz, J J: (1972) Semantic Theory. New York: Harper and Row.

Giordan H: (1975) L'enseignement de l'occitan. Langue Française No 25. pp 84-103.

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

Lafont, R: (1971) Un problème de culpabilité sociologique: la diglossie franco-occitane. Langue française. No 9. pp 93-9.

Lass, R: (1971) Boundaries as obstruents: Old English voicing assimilations and universal strength hierarchies. JL No 7, pp 15-30.

Mansion, J E: (1967) Harrap's Shorter French and English Dictionary. London: Harrap.

Martinet, A: (1965) De la morphonologie. La linguistique. No 1, pp 16-31.

Matthews, P H: (1974) Morphology: An introduction to the theory of word-structure. Cambridge University Press.

Saussure, F de: (1968) Cours de linguistique générale. Paris: Payot.

Schane, S A: (1968) French Phonology and Morphology. Cambridge, Mas.: MIT Press.

Schane, S A: (1971) The phoneme revisited. *Language*. Vol 47, pp 503-21.

(1972) How abstract is French Phonology? *In Generative Studies in Romance Languages*. Casagrande, J & B Saciuk (eds) Rowley, Mass.: Newbury House.

Séguy, J: (1951) *Le français parlé à Toulouse*. Toulouse: Editions Privat.

Warnant, L: (1968) *Dictionnaire de la prononciation française*. Editions J. Ducillot.

ON THE ROLE OF NOTATION AND THE ORDERING OF RULES IN PHONOLOGY

I M ROCA

The following morphophonemic alternations obtain in standard Eastern Catalan:

(1)

		Masc sing	Masc pl	Fem sing	Fem pl
/sant/	"saint"	[san]	[sans]	[sántə]	[sántəs]
/kamp/	"field"	[kam]	[kams]	[kəmpét]	[kəmpéts] (masc. diminut.)
/blaŋk/	"white"	[blaŋ]	[blaŋs]	[bláŋkə]	[bláŋkəs]
/alt/	"tall"	[al]	[als]	[áltə]	[áltəs]
/kurt/	"short"	[kur]	[kurs]	[kúrtə]	[kúrtəs]

(2)

		Masc sing	Masc pl	Fem sing	Fem pl
/san/	"healthy"	[sa]	[sans]	[sánə]	[sánəs]
/prim/	"thin"	[prim]	[prims]	[prímə]	[príməs]

In order to account for these data, the following two rules have been suggested (cf. Lleó, 1970: 29; Brasington, 1973: 25; Roca, 1975: 54, 56):

(3) $$\begin{bmatrix} +\text{coronal} \\ +\text{nasal} \end{bmatrix} \longrightarrow \emptyset \; / \; \_\_\_ \; \#$$

(4) $$\begin{bmatrix} +\text{obstruent} \\ -\text{continuant} \\ \alpha\text{anterior} \\ \beta\text{coronal} \end{bmatrix} \longrightarrow \emptyset \; / \; \begin{bmatrix} +\text{consonantal} \\ -\text{obstruent} \\ \alpha\text{anterior} \\ \beta\text{coronal} \end{bmatrix} \_\_\_ \begin{Bmatrix} \# \\ s \end{Bmatrix}$$

These two rules must stand in a counter-feeding relationship, as illustrated in (5):

```
(5)              / ≠   sant   ≠ /  ("saint")
     rule (3)          ------
     rule (4)        ≠  san  ≠
                       [san]       ("saint")
```

Derivations like that in (6) must therefore be avoided:

```
(6)              / ≠   sant   ≠ /  ("saint")
     rule (4)        ≠  san  ≠
     rule (3)        ≠  sa   ≠
                       [sa]        ("healthy")
```

Counterfeeding relationships are achieved within the standard model of generative phonology by means of extrinsic ordering. The principle of extrinsic ordering, however, has been challenged recently, and models of intrinsic ordering and free rule application have been put forward to replace it.

Kisseberth (1972) incorporates Kiparsky's diachronic simplicity metric into a synchronic model. Rules are thus ordered in such a way as to become "maximally transparent" (cf. Kiparsky, 1971: 623) where "a rule A→B/C__D is opaque to the extent that there are surface representations of the form (i) A in environment C ___ D, or (ii) B in environment other than C ___D".(Kiparsky, 1971: 621-622). In the case in hand, the counterfeeding ordering of (3), (4) will generate the opaque form [san], whereas the feeding order predicted by Kisseberth's model will yield the wrong surface representation for /sant/ ("saint").

The problem is posed by the actual existence of opaque rules, ie "rules that do not apply to certain phonetic structures meeting their structural description or that do apply even though the context which conditioned them is not present in the surface" (Kisseberth, 1972: 12).

The solution proposed by Kisseberth lies in the incorporation of global conditions into the general principle of minimisation of opacity. Along these lines, global condition (7) ought to be added to rule (3):

(7) provided that no segment between the nasal and the boundary has been deleted by the operation of another rule.

(7) will now block in derivation (6) the application of rule (3) after rule (4), which deletes the stop that follows the nasal and precedes the boundary in the underlying representation. Thus the combination of rule (3) and global condition (7) makes it possible to maintain intrinsic ordering for the Catalan rules (3), (4).

While Kisseberth's model does produce the correct output in the case under analysis, I shall not adopt it here. The reasons for this are twofold. First, the extent to which the use of global conditions differs from traditional extrinsic ordering is debatable. It seems quite clear that the incorporation of (7) into (3) is tantamount to imposing the condition that the application of (3) has priority over that of (4), ie to the introduction of an extrinsic ordering constraint on the application of these two rules. This is also the view of Dinnsen (1972: 14): "the derivational history specification serves no other purpose than extrinsic ordering would in the establishment of the counterfeeding relation. Given this, it is totally spurious to maintain that derivational history permits the abandonment of extrinsic ordering in any real sense". Kisseberth himself seems remarkably close to this opinion when asserting that "rule ordering is a way of encoding such facts [counterbleeding relationships, IMR] about the derivation ... Global conditions will generally be able to deal with cases where rule ordering is a device encoding derivational history" (1972: 18).

The other reason for the rejection of Kisseberth's model lies on the fact that neither the principle of minimisation of opacity nor global conditions can account for mutually bleeding relationships. Extrinsic ordering must thus be kept for this subset of rule relations, even if under the new name of "priority statements". As Dinnsen points out, such a model made up of the principle of minimisation of opacity complemented with global conditions would be interesting if it made the claim that the relationships it cannot define (ie mutually bleeding) do not exist in natural languages. This, however, is not the case, and extrinsic ordering cannot be dispensed with (cf.Dinnsen, 1972: 15).

Koutsoudas, Sanders and Noll (hereafter KSN) have recently argued against the existence of extrinsic rule ordering. Their basic principle is that "each rule simply applies to every representation that satisfies its structural description"(1974: 3). As must be obvious, this model cannot account for the counterfeeding relationship between rules (3), (4), since it will precisely predict the application of (3) to the representation obtained after the application of (4) to the underlying form.

We must now go back to the ordering statement made by the authors mentioned at the beginning of this paper, namely that rules (3), (4) stand in a relation of conjunctive ordering, the application of (3) taking precedence over that of (4). It must be recalled at this point that CH favour disjunctive over conjunctive ordering: "abbreviatory notations must be selected in such a way as to maximize disjunctive ordering" (1968: 63). Note that it is the claim of generative phonology that there exists a correlation between the formalism and the expression of true and significant linguistic generalisations, and that this is subject to empirical verification. Great emphasis is laid on this matter all through The Sound Pattern of English: the formal devices used for phonological description must "permit us to formulate general statements about the language which are true and significant, and must provide a basis for distinguishing these from other generalisations which are false, or which are true but not significant" (CH, 1968: 330); and "when we select a set of formal devices for the construction of grammars, we are, in fact, taking an important step toward a definition of the notion 'linguistically significant generalisation'. Since this notion has real empirical content, our particular characterisation of it may or may not be accurate as a proposed explanation" (ibid). Within this framework, it is crucial to determine the true relationship between the Catalan rules (3), (4).

(4) is in fact a schema standing for rules (8) to (11), which constitute its expansion:

(8) $$\begin{bmatrix} +\text{obstruent} \\ -\text{continuant} \\ +\text{anterior} \\ +\text{coronal} \end{bmatrix} \rightarrow \emptyset \; / \; \begin{bmatrix} +\text{consonantal} \\ -\text{obstruent} \\ +\text{anterior} \\ +\text{coronal} \end{bmatrix} \; \underline{\quad\quad} \; \begin{Bmatrix} \neq \\ \text{s} \end{Bmatrix}$$

(9) $$\begin{bmatrix} +\text{obstruent} \\ -\text{continuant} \\ +\text{anterior} \\ -\text{coronal} \end{bmatrix} \rightarrow \emptyset \; / \; \begin{bmatrix} +\text{consonantal} \\ -\text{obstruent} \\ +\text{anterior} \\ -\text{coronal} \end{bmatrix} \; \underline{\quad\quad} \; \begin{Bmatrix} \neq \\ \text{s} \end{Bmatrix}$$

(10) $$\begin{bmatrix} +\text{obstruent} \\ -\text{continuant} \\ -\text{anterior} \\ +\text{coronal} \end{bmatrix} \rightarrow \emptyset \; / \; \begin{bmatrix} +\text{consonantal} \\ -\text{obstruent} \\ -\text{anterior} \\ +\text{coronal} \end{bmatrix} \; \underline{\quad\quad} \; \begin{Bmatrix} \neq \\ \text{s} \end{Bmatrix}$$

(11) $$\begin{bmatrix} +\text{obstruent} \\ -\text{continuant} \\ -\text{anterior} \\ -\text{coronal} \end{bmatrix} \rightarrow \emptyset \; / \; \begin{bmatrix} +\text{consonantal} \\ -\text{obstruent} \\ -\text{anterior} \\ -\text{coronal} \end{bmatrix} \; \underline{\quad\quad} \; \begin{Bmatrix} \neq \\ \text{s} \end{Bmatrix}$$

These rules are ordered disjunctively in the sequence given above, according to CH's definition of the relation "precedes" (cf. 1968: 393).

Observe that rule (10) is not a rule of Catalan. This language, in effect, does not permit sequences of palatal plus stop, as formalised in the following negative morpheme structure condition:

(12) $\sim \begin{bmatrix} -\text{anterior} \\ +\text{coronal} \end{bmatrix} \quad [-\text{continuant}]$

Note, for example, that loans from Spanish comply with the statement expressed in (12):

(13)	Spanish		Catalan		
	rancho	[r̄aŋčo]	ranxo	[r̄aŋʃu]	"ranch"
	gancho	[ɡaŋčo]	ganxo	[ɡaŋʃu]	"hook"

The problem of vacuous subparts resulting from the expansion of a schema is not new. CH's defence of this type of rule lies in their alleged predictive power (cf. CH, 1968: 351). In the case in hand, this amounts to the statement that, although (10) is not included in the phonology of Catalan, nevertheless (10) is a possible rule to be brought about by language change in the specific case of Catalan, whereas the same rule with the values for the features "anterior" and "coronal" reversed in the environment is not a possible rule of that language. Interesting as this is, the justification for the inclusion of "possible rules" in the synchronic grammar of a language is by no means obvious. TG has been criticised for its excessive power, and incorporating non-existent rules in the grammar to permit the collapse of actual rules into general schemata does not do much towards restricting it. As has been said above, CH have put great emphasis on the task of the grammar as being the formulation of true and significant statements about the language. The extent to which the inclusion of ghost rules like (10) contributes to the establishment of true and significant generalisations about the language in question, however, seems debatable. The argument just put forward parallels that of McCawley when commenting on the effects of "possible rules" on the power of the grammar:

"One who takes "excessive power" arguments seriously has as his goal characterizing "phonological rule" so as to include all and only the phonological rules that the phenomena of a natural language could demand, and the mere fact that a certain putative rule would give correct answers if incorporated in a grammar of some language does not justify calling it a "possible rule" " (1971: 3). While McCawley is arguing at a more abstract level which bears on the form of the rules, it seems to me that his line of argument can be carried over to the case in hand. Needless to say, CH's claim of predictive power is an entirely empirical question. Should it be found empirically justified, the case of the inclusion of vacuous rules in the grammar would be strengthened, although the theoretical issue of evaluating synchronic simplicity against diachronic predictability would remain.

Given the discussion above, it can now be stated that schema (4) expresses a spurious generalisation on the Catalan data and must therefore be done away with.

The two problems to be solved at this stage are, on the one hand, the formalisation of the disjunctive ordering obtaining between rules (3) and (8), which would make them consistent with KSN's free ordering hypothesis, and, on the other, the collapse of rules (8), (9) and (11), the matrices of which contain identical features, the specifications of "anterior" and "coronal" alone being at variance. Rule (14) replicates rule (3) with a more fully specified matrix:

$$(14)\quad \begin{bmatrix} -\text{vocalic} \\ +\text{consonantal} \\ -\text{continuant} \\ +\text{anterior} \\ +\text{coronal} \\ +\text{nasal} \end{bmatrix} \longrightarrow \emptyset \quad / \quad \underline{\qquad} \; \#$$

Rule (15) is an alternative to (8), obtained by replacing the feature "obstruent" by "vocalic", "consonantal" and "nasal" in the segment to the left of the arrow (I have kept the feature "obstruent" in the environment as an abbreviatory device for convenience of exposition alone;

the possibility of replacing it by the three aforementioned features is obvious):

(15) $$\begin{bmatrix} -\text{vocalic} \\ +\text{consonantal} \\ -\text{continuant} \\ +\text{anterior} \\ +\text{coronal} \\ -\text{nasal} \end{bmatrix} \longrightarrow \emptyset \; / \; \begin{bmatrix} +\text{consonantal} \\ -\text{obstruent} \\ +\text{anterior} \\ +\text{coronal} \end{bmatrix} \_\_ \begin{Bmatrix} \# \\ s \end{Bmatrix}$$

(15) is a schema representing two conjunctively ordered rules, as follows:

(16) $$\begin{bmatrix} -\text{vocalic} \\ +\text{consonantal} \\ -\text{continuant} \\ +\text{anterior} \\ +\text{coronal} \\ -\text{nasal} \end{bmatrix} \longrightarrow \emptyset \; / \; \begin{bmatrix} +\text{consonantal} \\ -\text{obstruent} \\ +\text{anterior} \\ +\text{coronal} \end{bmatrix} \_\_ \#$$

(17) $$\begin{bmatrix} -\text{vocalic} \\ +\text{consonantal} \\ -\text{continuant} \\ +\text{anterior} \\ +\text{coronal} \\ -\text{nasal} \end{bmatrix} \longrightarrow \emptyset \; / \; \begin{bmatrix} +\text{consonantal} \\ -\text{obstruent} \\ +\text{anterior} \\ +\text{coronal} \end{bmatrix} \_\_ s$$

It is now a straightforward matter to collapse (14) and (16) following CH's conventions. In formal notation the schema will be as in (18):

(18) ($_{a1}$ +segment, +consonantal, -obstruent, +anterior, +coronal), +segment, -vocalic, +consonantal, -continuant, +anterior, +coronal, ($_{a2}$nasal), -segment, +WB, $\longrightarrow$ ($_{a3}$+segment, +consonantal, -obstruent, +anterior, +coronal), $\emptyset$, -segment, +WB:

$a_1 = a_3$ and $a_1 \neq a_2$.

More informally, (18) can be represented as (19):

(19) $$\begin{bmatrix} -\text{vocalic} \\ +\text{consonantal} \\ -\text{continuant} \\ +\text{anterior} \\ +\text{coronal} \\ -\text{anasal} \end{bmatrix} \longrightarrow \emptyset \; / \; \left(\begin{bmatrix} +\text{consonantal} \\ -\text{obstruent} \\ +\text{anterior} \\ +\text{coronal} \end{bmatrix} \right) \_\_\_ \#$$

The expansion of schema (19) is precisely rules (14), (16), which now stand in a disjunctive relationship.

Schema (8), now broken into schema (19) and rule (17), cannot be collapsed with schemas (9) and (11). The collapse of these two schemas, however, is straightforward:

(20) $$\begin{bmatrix}+\text{obstruent}\\-\text{continuant}\\\alpha\text{anterior}\\-\text{coronal}\end{bmatrix} \longrightarrow \emptyset \,/\, \begin{bmatrix}+\text{consonantal}\\-\text{obstruent}\\\alpha\text{anterior}\\-\text{coronal}\end{bmatrix} \_\_\_ \begin{Bmatrix}\#\\ s\end{Bmatrix}$$

All the processes illustrated in (1), (2), are now accounted for by rules (19), (17) and (20). While there is an obvious loss of simplicity, it could be argued that the new rules are well justified. On the one hand, rule (4) expresses a false generalisation, given away by the existence of morpheme structure condition (12). Also, the three new rules make clear the asymmetric relation obtaining in Catalan between the matrices in the structural descriptions of the three rules. Finally, the new rules do not call for any particular order in their application and are therefore consistent with the free ordering model proposed by KSN. While the evaluation of this model with respect to CH's standard model of extrinsic ordering is far from being a solved issue, the possibility of removing Catalan apparent counterexample of counter-feeding illustrates the crucial role of notation in generative phonology.

January 1975

ACKNOWLEDGEMENT

I am very grateful to Martin Atkinson for his careful perusal of the manuscript and for his mathematical illuminations.

REFERENCES

Brasington, R.W.P.: (1973)."Reciprocal rules in Catalan phonology". JL 9, 25-33.

Chomsky, N., and Halle, M.: (1968). The Sound Pattern of English. New York: Harper and Row.

Dinnsen, D.A.: (1972). "General constraints on phonological rules". Indiana University Linguistics Club.

Kiparsky, P.: (1971). "Historical linguistics". In Dingwall, W.O. (ed.), A survey of linguistic science. Maryland, University of Maryland.

Kisseberth, C.W.: (1972). "Is rule ordering necessary in phonology?". Mimeograph.

Koutsoudas, A., Sanders, G., and Noll, C.: (1974). "The application of phonological rules". Lg 50, 1-27.

Lleó, C.: (1970). Problems of Catalan phonology. Seattle: University of Washington.

McCawley, J.D.: (1971). 'On the role of notation in generative phonology'. Indiana University Linguistics Club.

Roca, I.M.: (1975). "Phonetics is not phonology". JP 3, 53-62.

EMG - TECHNIQUES AND APPLICATION TO SPEECH RESEARCH

Carol Riordan

The individual muscle fibre may be considered the structural unit of contraction. The amount the whole muscle can contract depends on the "maximum contraction of its contractile units, ie. its individual fibres" (Fromkin and Ladefoged 1966, 222). Due to the extensive branching of the axon, a single nerve innervates a number of muscle fibres. "The term 'motor unit' includes together with the muscle-fibres innervated by the unit, the whole axon of the motoneurone from its hillock in the perikaryon down to its terminals in the muscle" (Sherrington 1925, 519). The muscle comprises such motor units. The number of individual fibres may vary from muscle to muscle. Further, the "innervation ratio varies from 1:3 (one motor fibre innervates three muscle fibres) in extrinsic eye muscles to 1:150 - 1:200 in some leg muscles (Eyzaguirre 1969,40). Not all of a muscles' fibres are stimulated at once. Variations in movements - in their range, force and type - are ultimately determined by differences in the interaction and collaboration or motor units" (Brodal 1969, 123). A minimum load activates the most excitable fibres, ie those with the lowest threshold to stimulation. As the stimulus strength increases, fibres of lower excitability (higher threshold) are recruited. A further characteristic of this recruitment, "the unit which was first to appear in the contraction is the last to disappear: and vice versa, the unit which occurred last is the first to stop discharging" (Brodal, 1969, 124).

The basis of the excitation of muscle fibres is the biological membrane separating the intra- and extracellular fluids. Katz (1966, 42) describes the differing electrolyte content of the cell and of its external medium: "On the outside the major ionic constituents are sodium and chloride; inside the cell these ions amount to less that 15% of the electrolyte balance. Sodium is replaced by potassium which is accumulated to a concentration 20 to 50 times higher than in the external fluid". The resting state of the cell sees a certain balance maintained between the two substances. On the one hand, there exists an equimolarity of the inside with respect to the outside; that is, the number of particles on both sides of the membrane are approximately equal. A second property is electrical neutrality: the number of positively and negatively charged ions on either side of the membrane are about equal. The ionic composition of these solutions is the basis for certain electrophysiological phenomena. Specifically, "electrical potential differences, in general, appear at the boundaries between two electrolyte solutions if there are ions of different mobility or concentration on either side" (Katz, 1966, 48). In this case a selectively permeable membrane is placed across the boundary. This surface structure behaves as an ionic barrier to restrict ionic movement, ie independent mobility of ions. The derived potential difference is coloured by the presence/absence of such a chemically differentiating membrane. Further, the "sign and size of the p.d. (potential difference) across the membrane is determined by its relative permeability to the principal organic ions, sodium, potassium and chloride" (Katz, 49). The potential difference found in the resting cell is a diffusion potential. It is a secondary consequence of the "utilization of the metabolic energy supplied by the cell for the purpose of expelling sodium which has leaked into the cell and helping to accumulate potassium in the interior". The maintenance of this resting or membrane potential results from the fact that "the resting axon membrane has a much higher conductance for ... both potassium and chloride ... than for sodium" (Katz, 67-8).

In the active state, the NA^+ ions in abundance outside the cell take advantage of concentration and electrical gradients and permeate towards the cell interior. The regenerative entry of sodium effectively produces an all-or-none response once the threshold potential has been exceeded: "At this point the membrane passes through a state of unstable equilibrium" which may lead to a larger potential change, the action potential or spike. Once elicited, this "transient self-amplifying potential change ... propogates along the whole length of the fibre at constant velocity and without attentuation of signal strength. It leaves behind a short refractory period - a silent interval of one or a few milli-seconds during which the fibre is unable to carry a second signal. The system is ready to be re-excited and fire another propogated impulse" (Katz, 36).

Muscle cells, therefore, "respond to stimulation by irritating chemical processes which change the configuration of the cell" (Grossman 1967,8). Once stimulus strength is sufficient (the electrically unstable ignition point of the system) an action potential is produced. It is this "displacement of the resting membrane potential in the direction of depolarization which is invariably the stimulus which starts the chain of events leading to the development of tension and to muscle contraction" (Katz, 161). The stimuli for muscle activity are the consequence of a chemical transmission occuring at the neuromuscular synapse. A synapse is defined as a "functional contact between two excitable cells whose cytoplasms are enclosed within separate membranes (Katz, 97). In the case of neuromuscular junctions, the continuity of electrical transmission across the structurally isolated nerve and muscle fibres is established through the release by the nerve of a chemical transmitter in sufficient quantities to stimulate the muscle fibre. This "chemical 'mediator', acetylcholine, is secreted from the terminals. After being transferred across the synaptic cleft ... this substance produces the required local depolarization (and plate potential) of the muscle fibre" (Katz, 121). It is this depolarization which, if of sufficient amplitude (ie the threshold of the fibre), triggers the self-reinforcing action potential. Although the response of the individual fibres is an all-or-none affair, the contraction of the whole muscle can be graded. This is achieved, in part, though the synchronic neural excitation of a greater (or fewer) number of fibres, commensurate with the opposing load. The frequency of motor unit discharge is also an expression of stimulus strength (to the limits defined by the restoration factor of excitation). The operation of these two strategies - the number of available units innervated and their firing frequency - differs for varying degrees of voluntary effort. Fromkin and Ladefoged (1966, 223) report that at low levels of muscle action "the impulse rate of the individual active motor units is a good measure of the force being produced ...; but at higher levels, the degree of recruitment of additional motor units becomes a more important factor." The findings of Kugelberg and Skoglund (1946, 119) are cited in this reference:

> "The initial frequencies of the motor units associated with slight degrees of voluntary contraction range for the most part between 5 and 10/s, the frequency of the unit discharge increases with increasing contraction up to a maximum of about 30 - 50/s, the initial frequency for each new unit is, as a rule, lower than for those already involved, but with further increase of the contraction the new unit soon attains the same maximum frequency."

The electrical activity of muscles can be recorded both intra- and extracellularly. "Electromyography is the registration of muscle action potentials (which) give information as to the state of the muscle and indicate the activity of the motor neurons in reflex and voluntary contraction" (Buchtal 1957, 9).

A fibre diameter of 10-100μ in vertebrates allows for the insetion of a minute electrode into the muscle fibres in the case of intracellular recording. An electrode within a resting muscle fibre reveals a potential difference across the surface membrance of 90mV.

During excitation, the waveshape of the action potential is monophasic, and has an amplitude of the order of 80-120 mV (Eyzaguirre). In extracellular recording (to which the term EMG will refer in this paper) "the electrode picks up the action potential as it is conducted through the medium which surrounds the active fibre" (Buchtal, 9). Positioning of the electrodes outside the cell results in a biphasic action potential: the continuous movement of the signal source (impulse) causes a situation whereby the first electrode becomes negative and then positive with regard to the other. Further, withdrawal of the electrode from the fibre reduces the amplitude of the spike: "The impedance of the external medium is small as compared with the impedance of the fibre interior and hence the voltage of the extracellularly recorded potentials is maximally only 2 to 10% of the intracellularly recorded potential changes" (Buchtal: 10; see also Katz, 1966: 28-30).

The muscle action potential during voluntary contraction revealed by extracellularly recording "represents a summation of volume conducted activity from many muscle fibres belonging to the same motor unit" (Buchtal, 12). Both the duration and the maximum voltage of the potential resulting from this summation are greater than that of a single fibre. Buchtal (p.12) reports:

> "The greater duration is due to the fact that the points of impluse initiation for the different fibres are scattered throughout the above mentioned innervation zone (motor unit). The component of the motor unit potential which has the largest amplitude, the spike potential, arises from a small fibre group near the electrode ("sub-unit, 20-30 fibres), while the majority of fibres of the motor unit contribute to the more protracted initial and terminal parts of the potential".

Under conditions of weak contraction, the single motor unit potential can be discriminated. As maximum contraction is approached, however, potentials are recorded as an interference pattern of the activity of various fibres in the same motor unit. In this case, the number of different motor units being recorded simultaneously makes identification of the individual unit unreliable.

The duration of the single unit potentials may be defined as "the time interval between the initial deflection from the base line and the point at which the terminal deflection again returns to the baseline" (Buchtal, 15). The duration of potentials varies from muscle to muscle, but is the same for both increased and decreased contractile force. The amplitude of the recorded potential is influenced by many factors. In the first place, the structural characteristics of the muscle are important here: (1) All else being equal, the greater the number of fibres per motor unit, the larger the spike (Fromkin and Ladefoged, 226); (2) A fibre with a low resting potential will have a comparatively small action potential (Eyzaguirre, 41): and (3) "there is a discontinuity in the rate of decrease of the amplitude of the recorded spike when a muscle boundary or septum intervenes between the motor unit and the electrode; in these circumstances, the potentials recorded are of an order of magnitude smaller" (Fromkin and Ladefoged, 224). Further the "amplitude of the spike varies inversely with an increase of distance from the active fibres" (Ibid. 225). Finally, not only is the positioning of the electrode with respect to the active motor unit significant, but "it has been shown (Bauwens, 195; Buchtal, Guld, and Rosenfalck, 1957: Peterson and Kugelberg. 1949) that the size and type of recording device effects the duration and amplitude of the recorded action potential (Ibid; 226). The complexity introduced into the recorded potential due to the effect of neuronal, muscular and electrode transmission must be accounted for in any hypothesis as to the relation of electromyographic data to muscle activity or to higher systems of organization. The possibility of such an account at present will be developed shortly.

Motor unit potentials can be recorded from either surface of needle electrodes. The selection of one particular method depends, in part, on the location of the signal source. Insofar as each precludes the prescent of certain information in the myogram the nature of the sampled activity is a further consideration. The signal recorded by surface electrodes represents "the integrated potentials caused by the electrical disturbances in a large ... muscle mass" (Bauwens 1948: 136). The trace "appears as a continuous positive and negative variation is potential in which the action potential of individual units can no longer be distinguished" (Lenman 1969: 849). On the other hand, needle electrodes obtain "samples of electrical disturbances in a very restricted volume of the tissues" (Bauwens 1948: 136) wherein spikelike traces can be discriminated. Further, both mono- and bipolar recording is available.[1] Monopolar surface electromyography records action potentials with one of a pair of electrodes over active tissue, and the other over an inactive area. In the case of bipolar surface electrodes, both are positioned over active tissue. A typical example of each type of needle electrodes is described by Buchtal (p. 19):

1. Buchtal also describes multielectrodes and concentric needle electrodes which he has used. In the case of the former, six or seven different leads are contained at a distance of five to ten millimetres in a cannula to "facilitate simultaneous recording from a number of points at known distances from each other and give a well-defined measure of the spread of synchronous activity" (p. 19). Concentric needle electrodes reveal a potential measured between the end of an insulated wire and the surrounding cannula (as indifferent electrode).

"The bipolar electrode consists of two platinum wires insulated from each other and from the 0.65 mm steel cannula in which they are placed; the potential is measured between the two end surfaces of the platinum wires which have a distance of 0.5 mm between the centres of the leading-off areas.

"The unipolar electrode consists of an insulated metal needle; the potential is measured between the tip of the needle, from which the insulation is removed, and an indifferent plate electrode on the skin."

In bipolar recording, the receptive field is restricted by differential amplification, wherein signals common to both electrodes are cancelled. Consequently, "the greater the distance the signal source is from the electrodes, the greater the degree of cancellation, since differences in the amplitude and arrival time at each electrode will be minimized" (Hanson, Sussman, and MacNeilage: 1971, 2-3). Further, it is reported (p.4) that the "cancellation of low amplitude initial and final phases and, more importantly, on the directionality of the receptive field of the bipolar electrodes" produces a potential of about 50% shorter duration than the monopolar potential recorded simultaneously. Findings on the amplitude relationships from monopolar and bipolar derivations are reported in Mansell (1972, Ch. 6).

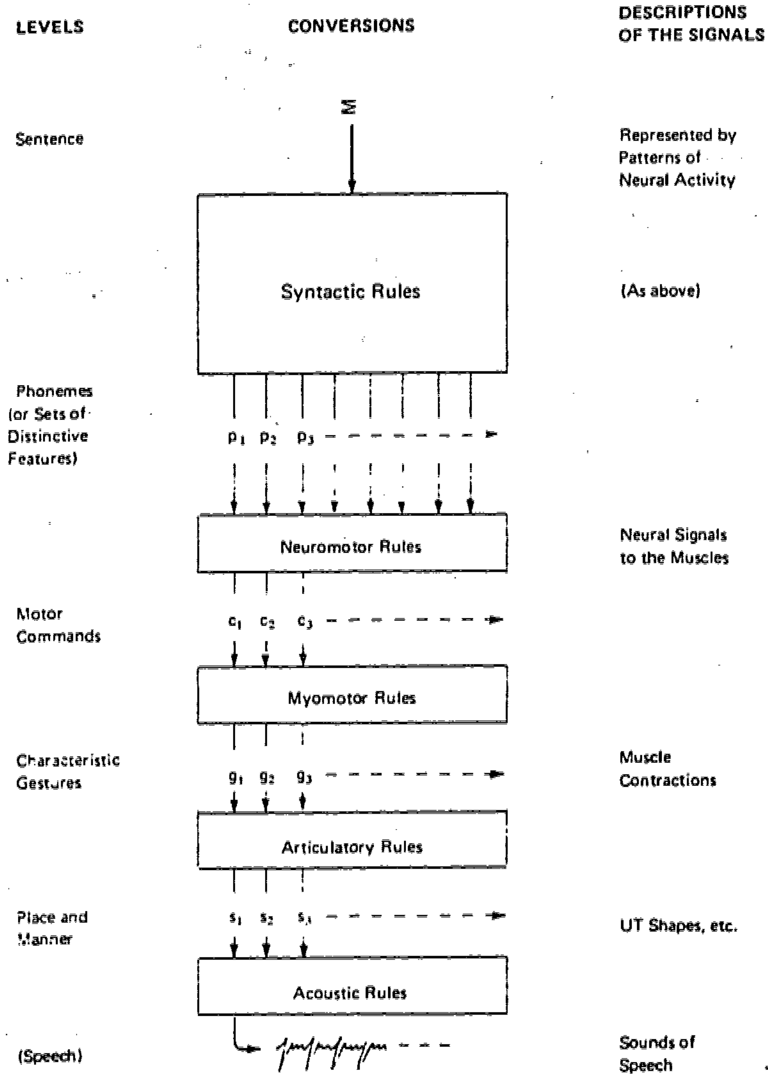
The original hypothesis as to the appropriateness of EMG for the study of speech production/perception includes an assumption as to the role of motor commands in articulation. Specifically, what is suggested is that the phenomena characteristic to the level of motor commands are "in striking contrast to what is found at the acoustic level"; namely, "the EMG potentials - and even more the motor commands inferred from them - bear a simpler relation to the perceived phonemes than does the acoustic signal" (Liberman et al 1967, 451). In other words, an invariance of motor commands was presumed - a hypothesis as to the extent to which there exist simple correlations between component gestures and phonemic units (Cooper 1965, 166). Implicit in such a hypothesis is an assumption as to the level and nature of the encoding of "phonological units (which) are discrete, invariant and timeless (into) articulatory and acoustic events which are continuous functions of time" (Mansell, 25). The "working hypothesis" of EMG put forth by Cooper (1965) and his colleagues at Haskins Laboratories (see Liberman et at 1962, 1965, 1967; Cooper et al 1958; Harris et al 1965; MacNeilage 1963) suggests that "a substantial part of the restructuring occurs below the level of commands"(Liberman et al 1967, 451), ie below the level of EMG. The viability of such a hypothesis can best be examined in the context of the model of speech production it implies. Liberman et al, 1967, will be the primary source for this exposition (see also the above references for a similar account). The schematic account of this model is presented here as Figure 1,(p.48).

The input to the phonological phase of speech production is assumed to be a syntactically-defined structure meeting all conditions for phonological interpretation (ie as would be determined in TG by the labelled bracketing of surface structure, lexical representations, and readjustment rules: Chomsky and Halle 1968). The form of this input structure is a string of phonemes, or, more specifically, the complex of features defining each of such segments.

The simplest of speech production models would assume that "sets of subphonemic features ... exist in the central nervous system as implicit instructions to separate and independent parts of the motor machinery" (p. 447). This is not to imply however, that even in the simplest of models all instructions are monotypic (p. 447):

> "These instructions might be of two types, "on-off" or "go to" ... In the one case, the affected muscle would contract or not with little regard for its current state (or the position of the articulator it moves); in the other, the instruction would operate via the γefferent system to determine the degree of contraction (hence, the final position of the articulator, whatever its initial position). Both types of

SCHEMA FOR PRODUCTION



instructions - appropriate, perhaps, to fast and slow gestures, respectively - may reasonably be included in the model.

Such a "sequence of neural commands corresponding to this multidimensional string of control instructions", may yet require certain "amplitude adjustments and temporal coordination ... in order to yield the neural impulses that go directly to the selected muscles of articulation and cause them to contract ... and supplementary neural signals to insure cooperative activity of the remainder of the articulatory apparatus" (p. 447). Such a conversion would occur at the level of Neurometer Rules. However, any reorganization of instructions which may occur at this level is limited. Indeed it is a necessary condition for this model that "there would be no reorganization of the commands to the "primary" actuators for the selected features ... (hence) the neural signs that emerge would bear still an essentially one-to-one correspondence with the several dimensions of the subphonemic structure" (p. 447).

The conversion from neural command to muscle contraction is mediated by Myomotor Rules (p.447).

> "If muscles contract in accordance with the signals sent to them then this conversion should be essentially trivial, and we should be able not only to observe the muscle contractions by looking at their EMG signals, but also infer the neural signals at the preceeding level".

That this implication does not receive any additional substantiation within the model will be commented upon later.

It is at the next level - the conversion of muscle contraction to vocal tract shape - that the occurence of the considerable restructuring (encoding) evident in the acoustic waveform is predicted. This complex relationship is determined by both spatial and temporal overlap of vocal tract configurations due to the structural and functional idiosyncracies of the articulatory apparatus (see Ohman 1967).

The final conversion is from "continuously changing shape to a modulated acoustic stream" (Liberman et al 1965, 1.12). This is essentially the predictable one-to-one correspondence on a moment-by-moment basis as described by Fant (1960).

The crucial issue for a speech production model is clearly the level at which phonemes are encoded (restructured) in the sound stream. In this model, the major encoding process occurs with the introduction of temporal and structural complexity, the so-called Articulatory Rules. Below that level, ie Acoustic Rules, no encoding appears involved. With reference to the encoding on levels preceeding the Articulatory Rules, Liberman notes(p. 448):

> "But what of the upstream conversion, particularly the one that lies between the neural representations of the phonemes and the commands to the articulatory muscles? We cannot at the present time observe these processes, nor can we directly measure their output - that is, the commands to the muscles. We can. however, observe some aspects of the contractions - for example the electromyographic correlates - and if we assume, as seems reasonable, that the conversion from command to contraction is straightforward, then we can quite safely infer the structure of the commands. By determining to what extent those inferred commands (if not the electromyographic signals themselves) are invariant with the phoneme, we can, then, discover how much of the encoding occurs in the conversion from contraction to shape (Articulatory Rules) and how much at higher levels".

Assuming such a "trivial" conversion from neural signals to contraction, the nature of these commands can then be predicted (p. 448):

> "For commands of the on-off type ... we would expect muscle contractions - and EMG potentials - to be roughly proportional to commands; hence the commands will be mirrored directly by the EMG potentials when these can be measured unambiguously for the muscles of interest.

"Commands of the "go to" type would presumably operate via the efferent system to produce only so much contraction of the muscle as is needed to achieve a target length. The contraction - and the resulting EMG signal - would then be different for different starting positions, that is, for the same phoneme in different contexts. Even so, the significant aspect of the command can be inferred, since presence versus absence and sequential position (if not relative timing) of the EMG signal persist despite even large changes in its magnitude".

Various EMG findings are reported as implicating the claim that "motor commands ... are more nearly invariant (with phonemic units) than are the acoustic signals" (p. 451).

"When two adjacent phonemes are produced by spatially separate groups of muscles, there are essentially invariant EMG tracing from the characteristic gesture for each phoneme, regardless of the identity of the other." (p.449).

"When the temporally overlapping gestures for successive phonemes involve more or less adjacent muscles that control the same structures, it is of course more difficult to discover whether there is invariance or not" (p.450)

"We ... find ... in the onsets and offsets of EMG activity in various muscles a segmentation like that of the several dimensions that constitute the phonemes".(p.450).

It should be noted that the extent to which the data itself does, in fact, support the model outlined above appears questionable. The validity of such an interpretation notwithstanding, however, it is the view propounded here that in many cases the strongest criticism which can be aimed at experimental evidence in its inappropriateness for the situation. In other words, a methodological criticism can be raised with regard to the realism of the data interpretation vis-à-vis the hypothesis being tested. It would appear more effective, however, to criticise the set of theoretical constructs upon which the model was based, ie. as to the extent to which it corresponds with something in nature.

A model is a bundle of hypotheses specifying the relationships between variables determining events. It is the specification of generalized relationships extracted from a restricted determination of significant face, from which predictions may result. To be more than just an academic exercise, a model need conform to a scientific method, offering a criteria for its evaluation. In other words, to refute a model on formal terms, ie that that offered specification of assumptions was not explicit or fully rigorous is trivial. Any model must be expected to have a degree of logical consistency, ie to specify sufficient restrictions so as to enable an explanation internally consistent with and justified by its goals. Thus, the criteria by means of which a model may be measured is not in terms of its formal statement, but the extent to which the hypotheses themselves may be tested. The more chances for a hypothesis to be disproven, ie the more tests it can be put to, the stronger that hypothesis is. Clearly, a hypothesis which cannot conceivably be refuted is vacuous, ie of no use or meaning in a scientific sense for the discipline which harbours it. Models offer predictions about both observable and unobservable events. In the case of the latter, theoretical constructs "enable us to deal with physical situations which we cannot directly experience through our senses, but with which we have contact indirectly and through inference" (Bridgeman 1927, 53). These constructs may be of two types: those admitting no other physical operations except those in its definition, and those that do (Ibid). The testing of a model offering predicitons about observable events differs from that of a model whose predictions are not verifiable by direct observation. In the former case, the measure of a theory is the extent to which these predictions conform to reality. In the case of unobservable phenomena, hoever, a model cannot be rejected by measuring its predictions with respect to factual statements. What need be tested - and, therefore, testable - are the original assumptions. Such a test of model would be to show that a construct of the first type (limited to only those physical operations in its definition) does in fact admit other operations.

Two objections of this nature can be raised with regard to the above model of speech production. In the first place, the validity of the constraint on higher level reorganization may be undermined by speech data implicating the presence of post phonological control elements. Secondly, it has been noted that two central assumptions of this model concern (1) the relation of sub-phonemic features to neural signals, and (2) the relation of these commands to the electromyogram. Specifically it is assumed that (1) the input and output of Neurometer Rules (ie phonetic component) correspond essentially in a one-to-one fashion: (2) that the Myomotor conversion is equally as simple: and (3) that the effects of the propogation of activity along the muscle, and of the EMG recording of this activity are, again, trivial is nature. It should be noted that these are Theoretical Constructs of the first type: these assumptions not only define operations at these levels, but imply the exclusive nature of these definitions. The extent to which such assumptions are realistic is clearly a crucial test for the model - one which, I would claim, cannot be definitively passed.

With regard to the first (linguistic) objection, the introduction of all complexity at the periphery (ie Articulatory Rules) implies a "trivial" phonetic component. Mansell (1972: 25) has labelled such accounts "peripheralist" since "they claim that peripheral factors in the motor system, combined with peripheral feed-back mechanisms can satisfactorily account for the conversion from discrete stored units to the flow of speech in its acticulatory and acoustic aspect. Hansell himself espouses an opposing "centralist" account, claiming the existence of a control mechanism mediating between a phon-ology and a motor system with a "power of prediction over peripheral factors" (p.27). Mansell offers support for such a centralist hypothesis with speech data - collectively referred to as articulatory variation which cannot be explained without recourse to such a control mechanism.[1]

The view propounded here is in agreement with Mansell's conclusion that the existence of a decision-making phonetic component is a necessary condition to account for various speech phenomena. Further, I would claim that such an account appears to be the more theoretically productive of the two, insofar as this solution may accommodate distinct speech data. Specifically, the goal of any linguistic theory must be not only to present a model of human linguistic behaviour "collectively", but also to account for the individual speaker's acoustic waveform (Kim, 1966). At some stage in the speech production process, therefore, "fixed table" values for linguistic units must give way to a set of rules handling the phonetic variation found in linguistically identical waveforms. What is implied, therefore, is a capacity to generate an infinite - yet bounded - number of "surface structures" from a single "deep structure". This surface form may vary only so far as to preserve the "meaning" of the deeper representation. Such transformations appear to parallel those which would be necessary given the effects of environmental pressures and neuromuscular tendency on a speech string. In this case, the intrinsic variability of the universal vocal mechanism is subject to linguistic demands. Hence, despite idiosyncratic differences (and the subsequent lack of absolute acoustic values corresponding to the phonemes of a language), the speakers of the same language are mutually comprehensible. The regulation of such idiosyncratically-determined variation appears to occur in much the same way as Mansell's non trivial phonetic component. It does not seem

1. It should be noted that this is not meant to imply that "the <u>total</u> neural and muscle activity will be in a one-to-one correspondence with the phoneme, but implicates only one or a few component parts, perhaps even the contraction of a single muscle in the extreme case" (p.1.18). Nonetheless, the underlying assumptions remain the same. Consequently, whether total activity or only a characteristic component is represented, their validity is crucial for the model.

unreasonable to suggest, therefore, that the regulation of such indiosyncratically-determined variation would involve a strategy similar to the interpretation process of Mansell's control mechanism, ie. a single device responsible for the control of all aspects of speech production which introduce linguistically insignificant variation into the speech string. Indeed, a certain transparency of explanation arises insofar as such a solution points to some unifying aspect of the speech production process.

Hence, with regard to the first point raised, there is no a priori reason to constrain the reorganization which occurs at the level of Neuromotor rules. Indeed, evidence to the contrary may be evoked to refute such a hypothesis. The second objection - the nature of the Myomotor Rules and the inferences which can be made about muscle activity from electromyograms - is clearly a more basic issue for the evaluation of the Haskins' Model. That is, should the validity of the data provided by EMG be questioned, the empirical basis of this and other accounts would be undermined.

In the conversion of neural signals into muscle contractions, the Haskins' Model claims that "signals map directly onto muscles and control their contractions ...: moreover, the muscular events are observable by electromyographic techniques." (Liberman et al 1965, 1.11). A homogeneity is assumed to exist, therefore, (1) between neural impulses and muscle response, and (2) between the initiated muscle activity and the myographic display.

The effects of the recording apparatus on the electromyogram (which clearly play a role in (2) above) have been briefly mentioned. Mansell (1972) surveys the literature in this regard and presents original experimentation to provide a detailed account of these considerations (see esp. Chapter 6). With regard to (1), the relationship of muscle activity and the nervous system is described by Bauwens (1950, 217):

"The subject (the muscle) acquires a new significance ... when it is realized that, except for glandular activity and some reactions of the blood vessels accompanying emotional disturbances, muscular activity is practically the only way by which the nervous system exteriorizes its hidden processes. Facial expression, twitching, nervous tension, and even screaming are examples of muscular activity which reflect the state of the nervous system. The fact that the behaviour of muscles is so intimately related to this system allows the clinician to speak of them as its mouthpieces, in the same manner as he speaks of the eyes as its windows."

Such a special relationship of gross muscular and nervous activity, however, does not imply an isomorphism with regard to the units of each. Partridge and Huber (1967; 1276-1277) note the problems of EMG interpretation, its relation to motor control notwithstanding:

"From EMG it is possible to determine the participation of a specific muscle in a particular movement ... Nevertheless, it is not yet possible to examine an EMG record and interpret any detail of the movement pattern resulting from that neural-muscular activity".

The reason for this difficulty is, at least in part, a heuristic one (p. 1277):

"The samples activity of one unit may not be parallel to the activity of other unsampled units in other parts of the same or synergistic muscle and does not account at all for the action of antagonist muscles".

Beyond this, however, lies the difficulty of the undefined relationship of muscle input and output. Knowledge is particularly limited with respect to the "response to the dynamic signals essential to movement control" (Ibid). Their experimental findings indicate the complexity of the input-output relationships in the muscle insofar as pulse-rate alone does not determine muscle position (p. 1287):

> "In order to infer movement from pulse rate information under controlled experimental conditions, three factors had to be taken into consideration. These are input-output non-linearity; response lag: and hysteretic response of muscle. To infer physiological movement from EMG, it would seem that these same three factors and also the role of synergistic and antagonistic motor units, gravity, etc., would have to be considered".

Furthermore, the shape on an action potential is also partially determined by neuronal transmission factors (Hanson, Sussman, and MacNeilage 1971). The signal propogated along the length of the axon will vary little in amplitude and shape. The velocity of the wave, however, is closely related to the size of the fibre (Katz: 91-92). Consequently, due to the "differing lengths and diameters of the axon terminal collaterals, the neuronal impulse will reach the myoneural junctions of the motor unit with some asynchronism introduced" (Hanson et al; 1). This may be significant insofar as these factors appear to be of increased importance in the case of the particular fibre architecture found in much of the speech musculature (Ibid).

It appears therefore, that the assumption of homogenity from subphonemic features to an EMG record cannot be definitively confirmed by the available literature. The alternate hypothesis is likewise meaningless, in the sense that its validity is equally as precarious. A model which does not require this assumption therefore would appear a more meaningful and productive means at present to explain the speech production process. I would suggest that a claim for a decision-making component would determine such a model. Mansell notes that a contralist model presupposes the triviality of peripheral reorganization: "if the centre is capable of predicting behaviour of the periphery ... then it is capable of predicting the complexity introduced at the periphery and of, if necessary, modifying the instructions issued to the periphery on the basis of this knowledge" (p. 27). In other words, "to predict the future (of an event) is to carry out a certain operation on its past" (Wiener, N: 1961, 6). Similarly, further, whatever the transfer function of neuronal, muscular and electrode transmission may be, it could be expected that this information would be embodied in a non-trivial phonetic component and influence its decision-making.

Clearly, the stronger the assumptions underlying a theory, the stronger the predictions of that theory, ie the weaker its descriptive generality. Thus, often within a model, the more restrictive the nature of the assumptions, the more precise its predictions. However, in comparing two theories, with the same predictive power, the theory with the less restrictive assumptions is a better theory in a scientific sense in that it incorporates and extends beyond the other. As Mansell suggests, one effect of a hypothesis introducing all significant complexity at the periphery is "to divorce the central intention from the articulatory output (which) must necessarily entail the existence of a highly specific peripheral feedback system for the automatic maintenance of any structure in the articulatory output" (p.27). The claim of a phonetic component insofar as the involvement of such a peripheral device to account for the precision of controlled movement in speech can be inferred but not proven, a model avoiding such an assumption is a stronger model. Clearly, should the role of such a time-sharing strategy of the central nervous system in speech be established, the reverse would be the case.

Further, I would suggest that the "trivial" Myomotor conversion assumed in the Haskins model is not a necessary condition of a centralist approach per se. In this case, it is the appropriateness of EMG which is at stake. Yet, it should be noted that the viability of electromyography as a research technique, and the validity of a speech production model are independent issues. A dependence of the latter on the former is created only when EMG data constitutes supporting evidence, in which case the model is open to criticism from a methodological standpoint.

The predictions of a model can be no greater than the demands of the general philosophical considerations underlying it. The predictive expression of a model therefore, is presupposed by the assumptions and goals for which it was formulated. In other words, the very restrictions labelled as "necessary conditions" for the Haskins model constrain its scope.

There is no a priori reason to assume that the encoding of the speech stream is a language universal phenomena, ie occurs due to the characteristics of the vocal apparatus shared by all speakers. The constraint this prediction of a restructuring ignorant of any language-specific demands imposes, is the motivation par excellence for a so-called centralist model; namely, to formulate a testable model capable of accommodating a greater proportion of the available empirical data.

REFERENCES

Bauwens, P. (1948) Electromyography. British Journal of Physiological Medicine, 11, 130 - 136.

Bauwens, P. (1950) Analysis of action potentials in EMG. Proceedings of the Institute of Electrical Engineers, 97, 1: 217 - 221.

Bridgeman, P. (1927) The Logic of Modern Physics. MacMillan: New York.

Brodal, A. (1969) Neurological Anatomy. Oxford University Press: New York.

Buchtal, F. (1957) An Introduction to Electromyography. Copenhagen.

Buchtal, F., Guld, C., & Rosenfalck, P. (1957) Multielectrode study of the territory of a motor unit. Acta Physiologica Scandanavica, 39, 1: 83 - 104.

Chomsky, N. & Halle, M. (1968) The Sound Pattern of English. Harper & Row: New York.

Cooper, F. S. (1965) Research Techniques and Instrumentation: EMG. American Speech and Hearing Reports, 1, 153 - 168.

Cooper, F. S., Liberman, A. M., Harris, K. S., & Grubb, P. M. (1958) Some input - output relations observed in experiments on the perception of speech. Proceedings of the Second International Congress on Cybernetics, Namur, Belgium. Association Internationale de Cybernétique.

Fant, G. (1960) Acoustic Theory of Speech Production. Mouton: The Hague.

Fromkin, V. & Ladefoged, P. (1966) Electromyography in speech research. Phonetica, 15, 219 - 242.

Grossman, R. C. (1967) Methods of determining oral tactile experience. In J. F. Bosma (Ed.) Symposium on Oral Sensation and Perception. Charles C. Thomas: Springfield, Illinois.

Hanson, R. H., Sussman, H. M., & Macneilage, P. F. (1971) Single motor unit potentials in speech musculature. Paper presented at the VIIth International Congress of Phonetic Sciences, Montreal, August, 1971.

Harris, K. S., Lysaught, G. F., & Schvey, M. (1965) Some aspects of the production of oral and nasal labial stops. Language and Speech, 8, 135 - 147.

Katz, B. (1966) Nerve, Muscle and Synapse. McGraw-Hill: New York.

Kim, C-W. (1966) The linguistic specification of speech. UCLA Working Papers in Phonetics, 5.

Kugelberg, E. & Skoglund, C. R. (1946) Natural and artificial activation of motor units - a comparison. Journal of Neurophysiology, 9, 339 - 412.

Lenman, G. (1969) Integration and analysis of the EMG. In Walton (Ed.), Disorders of Voluntary Muscles. Churchill.

Liberman, A. M., Cooper, F. S., Harris, K. S. & MacNeilage, P. F. (1962) A motor theory of speech perception. Proceedings of the Speech Communication Seminar, Volume 2. Stockholm, Royal Institute of Technology.

Liberman, A. M., Cooper, F. S., Harris, K. S., & MacNeilage, P. F. (1965) Some observations on a model for speech perception. Haskins Laboratories SR-1, 2.1 - 2.27.

Liberman, A. M., Cooper, F. S., Shankweiler, D. P. & Studdert-Kennedy, M. (1967) Perception of the speech code. Psychological Review, 74, 431 - 467.

Mansell, P. (1972) Theoretical and Experimental Approaches to a Model of Speech Production. PhD thesis, University of Essex.

MacNeilage, P. F. (1963) Electromyographic and acoustic study of the production of certain final clusters. Journal of the Acoustical Society of America, 35, 4: 461 - 463.

Ohman, S. E. G. (1967) Numerical model of coarticulation. Journal of the Acoustical Society of America, 41, 2: 310 - 320.

Partridge, L. D. & Huber, F. C. (1967) Factors in the interpretation of the electromyogram based on muscle response to dynamic nerve signals. American Journal of Physical Medicine, 46, 1276 - 1285.

Peterson, G. E. & Kugelberg, E. (1949) Duration and form of action potential in the normal human muscle. Journal of Neurological and Neurosurgical Psychology, 12, 124 - 125

Sherrington, C. S. (1925) Remarks on some aspects of reflex inhibition. Proceedings of the Royal Society, B, 97, 519 - 545.

Wiener, N. (1961) Cybernetics; or control and communication in the animal and the machine. 2nd Ed. MIT Press, Cambridge, Mass.

SOME OBSERVATIONS ON THE ROLE OF PLACE AND MANNER OF ARTICULATION IN THE PERCEPTION OF THE VOICING CHARACTERISTICS OF FINAL CONSONANTS

Donal O'Kane

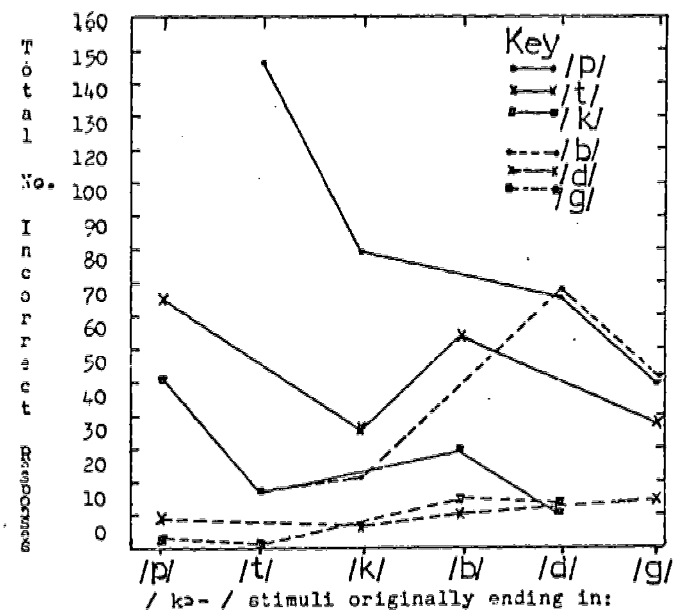
INTRODUCTION

In an experimental investigation recently conducted by the author of certain perceptual cues posited as determining the voicing status of postvocalic English stop consonants (results reported in O'Kane (forthcoming)) subjects were required to listen to three separate series of CV-tokens taken from the following monosyllables: /kɔp, kɔb; kɔt, kɔd; kɔk and kɔg/. Each individual series presented consisted of tokens made entirely from

(1) One of the above homorganic pairs
(2) Hybrid CV- syllables made by simply switching the vowels between the members within each homorganic pair. Finally, the vowel in each natural and each hybrid CV- syllable was progressively cut back in 20ms steps to within 14 - 20 ms of onset, five tokens being recorded at each duration. The tokens thus prepared were randomly ordered on magnetic tape and presented via a loudspeaker in three separate sessions (Session 1: /kɔp, kɔb/ and their hybrids; Session 2: /kɔt, kɔd/ and their hybrids and Session 3: /kɔk, kɔg/ and hybrids) to twenty subjects.

A forced choice was imposed on subjects' responses. The following example will suffice to illustrate. In session 1, subjects were instructed to decide whether a /p/ or a /b/ followed each CV- syllable they heard. Because subjects tend, in certain cases - mainly as a function of changing CV-duration - to think they may have heard or to subjectively report "hearing" other stop

FIGURE 1.



Total number of incorrect / p, t, k; b, d, g / responses (except for homorganic responses) made to each stimulus type.

consonants than /p/ or /b/, they were instructed in those circumstances to give these "incorrect" responses in place of /p/ or /b/. The same procedure applied to the other homorganic pairs. Thus, overall, the possible responses were /p,t,k; b,d,g/ in each session though instructions biased the subject towards "correct" responses. In all, subjects provided 9,000 responses and the experimental design enabled the categorisation of incorrect responses in terms of place and manner of articulation. Over half the total responses in each session consisted of incorrect responses - despite subjects' prior knowledge that each incorrect response was the result of a perceptual (linguistic) illusion. This bias, provided by prior knowledge, in fact provides more weight to the observations which will be made here.[1]

PLACE AND MANNER FACTORS

By considering the incorrect responses only, it was hoped to establish whether any general response tendencies could be found to indicate the involvement of place or manner of articulation characteristics which might distinguish certain groupings among the stop consonants studied. Table 1 provides the overall relevant data breakdown and Figure 1 presents the total number of incorrect /p,t,k,b,d,g/ responses - except for homorganic responses - made to the six different stimuli types used in the experiment.

1. All observations are based on responses to normal (i.e. non-hybrid) tokens.

TABLE ONE						
Total Number/Type of Responses Made to Each Stimulus Independently of Duration. *Round brackets contain deleted segments.						
	RESPONSE GIVEN					
	/p/	/t/	/k/	/b/	/d/	/g/
STIMULUS /kɔ(p)*/	399	75	51	63	9	3
/kɔ(t)/	144	389	17	17	32	1
/kɔ(k)/	85	36	436	23	7	113
/kɔ(b)/	276	64	30	406	11	13
/kɔ(d)/	77	333	11	78	488	13
/kɔ(g)/	51	38	352	52	15	392

CV- syllables cut from voiced-final tokens tended as a group to elicit responses in terms of manner of articulation. If we consider voicing status, incorrect responses display a strong contrast between the originally /b,d,g/-ended and the originally /p,t,k/-ended CV- syllables. Incorrect responses to originally /b,d,g/-ended stimuli reveal that subjects prefer voiceless responses: the ratio of voiceless to voiced responses in all homorganic pairs within the /b,d,g/-ended set being significantly greater than the voiceless-voiced ratio between pairs in the /p,t,k/-ended set. Tables 2 and 3 illustrate this.

TABLE TWO		
Number of Incorrect Voiceless and Voiced Responses given to Stimuli Originally Ending in /p,t,k,b,d,g/.		
Stimulus,	Number Incorrect Voiceless Responses	Number Incorrect Voiced Responses
Originally Ending in /p/	126	75
/t/	161	48
/k/	121	141
/b/	370	24
/d/	421	91
/g/	441	67

TABLE THREE		
Correct Responses to the Stimuli /p,t,k,b,d,g/ Seen Against Incorrect Homorganic Responses to the Same Stimuli.		
Stimulus Originally Ending in	Total No. Correct Responses	Total No. Incorrect Homorganic Responses
/p/	399	63
/t/	389	32
/k/	436	113
/b/	406	276
/d/	488	333
/g/	392	352

In contrast, voiceless tokens do not present such a clear picture. For CV- syllables originally ending in /p/ or /t/ the majority of incorrect responses are in terms of a place of articulation change with /p/ eliciting a majority of /t/ responses and /t/ a majority of /p/ responses. In general, then, when uncertainty (caused by vowel shortening) exists as to the voicing status of a following stop, the tendency - given a CV- stimulus taken from a /b,d,g/-ended monosyllable - is to make a homorganic voiceless response. Given the same uncertainty when presented with originally /t,p/-ended stimuli, the tendency is to choose a voiceless response with a differing place of articulation. The behaviour of originally /k/-ended stimuli in eliciting a majority of incorrect /g/ responses must remain unexplained here: Cf. Wajskop and Sweerts (1973).

If we regard, for the moment, the behaviour of originally /k/-ended tokens as peculiar, we have a situation like that shown in Table 4 where, overall, the most generalizable incorrect responses seem to be voiceless responses as opposed to voiced responses for which quite specific voicing cues seem necessary, lack of which result in a generalized voiceless response.

TABLE FOUR			
Total Number of Times /p,t,k,b,d,g/ Were Given as Incorrect Responses			
/p/	633	/b/	233
/t/	546	/d/	74
/k/	461	/g/	143*
* 30 if responses to /k/ are omitted.			

With regard to place of articulation, the further back in the oral cavity one goes, the less generalizable the cues for voicing in postvocalic, final stop consonants become.

REFERENCES

O'Kane, Donal: (forthcoming) "Manner of Vowel Termination as a Perceptual Cue to the Voicing Status of Postvocalic Stop Consonants".

Wajskop, M and Sweerts, J: (1973) "Voicing Cues in Oral Stop Consonants" Journal of Phonetics 1, 121-130.

OROSENSORY FEEDBACK MECHANISMS AND SPEECH PRODUCTION

by

Donal O'Kane

Department of Language and Linguistics

University of Essex

1.0 The coordination of nervous and muscular activity which characterizes speech production depends on a very close interrelationship between efferent and afferent neural systems. If human bodily afferent activity is lost, movement is essentially abolished (Henneman (1968b)). Hardy (1970) states that these two neurological systems are virtually impossible to separate physiologically and their interaction should therefore be considered in terms of a concept of " sensorimotor systems "(p50). This introductory section will be concerned with presenting a simple outline of the various orosensory receptor systems relevant to those aspects of speech production which are the subject of this paper.

1.1 The orofacial complex is relatively densely innervated with sensory receptors for relaying tactile, pressure and stretch information to higher levels. There are three major afferent paths relevant to speech production. These are: the dorsal root ganglia, the tractus solitarius and the trigeminal complex.

Current evidence seems to show that afferent information from the tongue enters the CNS via the dorsal root ganglia (Konigsmark (1970), p11). Bowman and Combs (1969) showed that lingual spindle afferents are present in the distal hypoglossal nerve and pass via the ipsilateral dorsal roots to the thalamus and then project to the lingual sensorimotor areas in both the pre- and post-central gyrii.

The tractus solitarius receives sensory information from the larynx and possibly part of the face and relays it to the CNS. It also receives input from the spinal cord and cortex and has muscle fibre connections with the cerebellum.

The sensory system comprising the trigeminal complex has the most discrete projections into the cerebral cortex and thalamus. Discrete cell clusters within this complex will respond to stimulation of the oral area " with specific cells responding to very localized stimulation of specific intraoral structures "(Jerge(1970),p75).Movements of the tongue and tongue shape are principally controlled by reflex mechanisms activated by trigeminal afferents (Kawamura(1970)).The trigeminal complex contains four major nuclei:the trigeminal ganglion,the main sensory nucleus of the Vth. cranial nerve,the nucleus supratrigeminalis and the mesencephalic nucleus of the Vth. cranial nerve.The receptive fields of the trigeminal ganglion are affected by stimulation of the mandible skin field,tongue,teeth and jaw and by thermal and tactile stimulation of the tongue.Information about light tactile stimulation is known to go from here to the thalamus(Konigsmark(1970),p13). The main sensory nucleus has a considerable projection area from the trigeminal nerve and is activated by light touch or pressure on the mouth or face(Konigsmark(1970),p14).The third major nucleus of the trigeminal complex - the nucleus supratrigeminalis - is primarily responsible for mandible movement and pressure on oral structures,specifically the teeth,palate and tongue.It is probably active in jaw reflexes and jaw movements in speech as it is located near the mesencephalic main sensory and motor nuclei of the Vth. nerve.It is also known to be directly connected with the cerebral cortex(Jerge(1970). Finally,the mesencephalic nucleus conducts proprioceptive information from the hard palate and teeth and has primary ganglion cells in the CNS.Jerge (1970) found it to contain two types of neurons:

(a) those innervating muscle spindles in the masseter,temporalis and medial pterygoid muscles and
(b) those innervating dental pressure receptors in the teeth.These neurons are probably active in jaw movements.

Secondary touch cells are also present with small oral receptive fields and these send axons to the trigeminal tract.

1.2 The principal CNS coordinating mechanisms which receive and either relay this information to the motor centres or modify outgoing efferent impulses according to the afferent input are: the thalamus,the cerebellum and the sensory cortex.

Various nuclei in the thalamus provide the final relay for all orosensory input.For instance,the ventral posterior medial nucleus relays facial sensation and Mountcastle and Henneman(1952) found stimulation of the intraoral structures,lips and larynx represented there.The thalamus relays this information via projections to the sensory cortex.

The cerebellum is a major centre for coordinating afferent input.For example,the hypoglossal nucleus which innervates the tongue muscles projects into the cerebellum as do afferent fibres from the tractus solitarius.The cerebellum works closely with the primary motor cortex by distributing impulses from the latter and by supplying it with afferent information. Its role is unconscious. As Henneman says: "it receives a continuous stream of impulses from receptors in muscles, joints, tendons and skin.... These sensory impulses do not mediate conscious sensations, but supply sensory cues essential to control of movement"(p1771). Damage to afferent inputs to the cerebellum or to the cerebellar mechanisms are known to cause speech defects(Kaplan(1960)). It is somewhat similar in function to the reticular formation which lies beside the trigeminal motor nucleus and coordinates a great variety and type of motor and sensory inputs at the brainstem level.

The sensory cortex is situated in the post-central gyrus and afferent information arriving there arises in conscious sensation(Hardy(1970), p53). The lower third of the gyrus (the sensory cortex area) receives information from the face, jaw, lips, tongue and pharynx. The largest part of this area is taken up by afferent activity from the tongue and lips. There is an abundance of association fibres which connect the post- and pre-central gyrii, the latter being in a " recurrent circuit [with(DO'K)] the basal ganglia and thalamus "(Konigsmark(1970), p6).

To sum up: the various sensory receptors of the orofacial complex are activated by tactile, pressure and proprioceptive stimulation and send this information via the dorsal root ganglia, the tractus solitarius and the trigeminal complex to the thalamus, cerebellum and, most importantly, the post-central gyrus. These sensory nuclei can also have their activity dampened or enhanced when necessary by the reticular formation, cerebellum or pre-central gyrus. This information is coordinated with efferent impulses from the primary motor cortex and, utilizing the smoothing and integrating capabilities of the cerebellum and basal ganglia - which have projections from the pre-central gyrus - the impulses project via the major motor nuclei - hypoglossus, nucleus ambiguus, facial nucleus and motor nucleus of the Vth. nerve - to the individual muscles of the orofacial complex involved in speech.

Finally, as was previously mentioned, the post-central gyrus, which is the principal sensory reception area of the cortex, uses a third of this area for receptive fields of the face and intra-oral area - with the tongue having relatively the largest representation(Jerge(1970), p77). Jerge notes that " the number of cells innervating a unit area of the oral cavity is much greater than for other parts of the face"(p66). It has also been found that the ability of an area of the orofacial complex to discriminate sensory stimuli varies " directly with the size of its cortical and thalamic projection areas. The relative volume of tissue of thalamic relay nuclei and of the post-central cortical gyrus is related directly to the density of peripheral neural innervation of that region and inversely related to the size of the receptive fields contained in this peripheral area"(Ringel and Fletcher(1965), p395).

2.0 It is not enough,of course,merely to establish the presence of sensory receptors in the oral cavity. Effective evaluation of the relative sensitivity of speech articulators to sensory stimulation is also necessary and requires the use of relevant tests.The present section will deal with one of the most effective recent approaches to this problem. Normal speech production requires accurate positioning and movement of articulators according to a relatively strict time schedule and is,to an extent,dependent on intact orosensory tactile receptors which relay accurate sensory information to the brain. Disruption of these input channels can result in speech output disorders(Ringel et al(1970);McDonald & Aungst(1967); Bishop et al(1973)).In recent years,the development of tests of oral stereognosis - defined as " the ability to recognize the form of objects through the sense of touch "(Ringel et al(1970),p410).have provided normative data on oral perception processes which are basic to all studies of orosensory feedback in speech(Locke(1968);Ringel et al(1965);Ringel and Ewanowski(1965);Rosenbek and Wertz(1973);McDonald and Aungst(1967,1970);Shelton and Hetherington(1967). Other tests - of two point discrimination,texture discrimination etc.(Ringel and Ewanowski(1965); Ringel and Fletcher(1967) - have also shown a correlation to exist between articulatory proficiency and oral perception but the oral stereognosis test is the most significant (with regard to speech production) and the best documented.

2.1 The oral stereognosis test now accepted as most consistent and reliable is the test used by Ringel et al(1968,1970) which is designed to test the subject's ability to discriminate between orally presented shapes.For an oral stereognosis test, one control and one or more experimental groups are usually required.The control group must have no history of motor or sensory disturbances and the experimental groups are usually (but not necessarily) articulatory defectives,grouped according to their speech defects - these defects being ranked in order of severity after examination by experienced speech pathologists.The stimuli used for the test are taken from a set of twenty geometrical forms (of heat resistant plastic) developed by the U.S. National Institute of Dental Research.The items chosen represent a wide range of difficulty and confusability which extensive testing has established(McDonald and Aungst(1967,1970);Ringel et al(1968). Because better,more reliable estimates of relative articulatory proficiency can be obtained using results from " between-class " comparisons(Rosenbek and Wertz (1973);Ringel et al(1970)) the forms used are further divided into four groups for later statistical analyses:oval,triangular,bioconcave and rectangular. Each pair is used once only (10 random pairs being later re-evaluated for reliability) and so every subject is presented with 65 stimulus pairs in all. The subject does not see the stimulaus form or touch it manually as this procedure can introduce intersensory artefacts.The first form is placed in the subject's mouth for five seconds,and after a five second interval,the comparison form is inserted for the same period.The presentation is randomized.While in the mouth the forms can be manipulated as desired by the subject.The importance of these tests is that they provide a method for judging the functioning of the speech mechanism and,as a result, can provide normative data against which to judge the degree of relative speech efficiency which differentiates speakers.The efficiency of orosensory function being a matter of degree rather than a matter of absolute dichotomy between normal sensory function and no sensory function,the finding that statistical analysis of between-class error rates can differentiate between levels of sensory dysfunction is of paramount importance(Rosenbek and Wertz(1973)).

2.2 Studies of articulatory defective speakers show that those subjects with articulatory defects have higher error scores than normal speakers and the errors made tend to increase as a function of the severity of the articulatory defect(Ringel(1970); Bishop et al(1973)).Rosenbek and Wertz(1973) found significant differences between apraxic patients and aphasic patients without apraxia,and further,that " the more severely apraxic the patient, the more poorly he performs " on oral stereognosis tests(p32).Stereognosis has now been recognized as an important evaluator of the integrity of the CNS and such tests as those under discussion are now generally used in neurological examinations(Ringel et al(1970).

This ability to identify forms orally increases with age until middle adolescence and remains relatively constant until old age when it can deteriorate(McDonald and Aungst(1967),p70).Children generally err more than normal or mildly articulatory defective adults.This is generally attributed to maturational factors,for example,oral cavity size and the improved motor abilities of adults(Ringel et al(1970)).There is also evidence which suggests that as the maturation process continues,a child will change from extensive dependence on auditory feedback to a greater dependence on tactile or kinesthetic feedback or both(Shelton and Hetherington(1967)).

2.3 The question of whether or not the degree of orosensory dysfunction in a speaker is the result of a motor deficit or a sensory input deficiency is of obvious importance.Earlier studies of oral stereognosis suggested that poor performance on the tests was a function purely of poor motor ability alone (McDonald and Aungst(1967)).Obviously,as Locke(1968) points out,oral stereognosis involves both the peripheral tactile receptors, " central integrating processes,and a minimum level of motor facility" (p1259).

It would not seem to be simply a higher level general perceptual ability as McDonald and Aungst would suggest.Edwards(1970) showed that speakers with articulatory defects who had higher error scores than their normal counterparts in oral stereognosis tests,did not produce significantly different error scores than normal speakers when the forms were manually explored.Rosenbek and Wertz (1973) demonstrated that apraxia of speech is not simply a motor (speech output) disorder but also a sensory (speech input) disorder.Apraxic patients could be differentiated on the basis of their orosensory perceptual ability.This deficiency is related,they go on, " to ability to process,store, recall and compare sensory and motor information received from the oral periphery "(p258) and disruption of orosensory feedback systems can result in speech output disorders.

2.4. Data from oral stereognosis and other sensory tests can help establish the relative sensitivity of various parts of the oral cavity.This data becomes more significant if considered in the light of evidence of oral cavity innervation presented earlier.As was noted in the previous section,various parts of the oral cavity are more densely innervated than others.Generally speaking,the density of oral cavity innervation(per unit area)decreases as we move from the anterior to the posterior part of the oral cavity.Ringel and Ewanowski(1965) examined oral sensitivity using a test of two point discrimination and found that the tongue tip had the greatest tactile sensitivity and ,in decreasing order of sensitivity,were the fingertip,upper lip, soft palate,alveolar ridge and thenar eminence of the right hand.Except for the fingertip,all tongue tip sites studied (at the midline and to the right and left of the midline) were significantly more discriminating than all other structures studied(p395).In a series of texture discrimination tests Ringel and Fletcher(1967) found that the lingual structures were , in general , " more accurate evaluators of texture " than other oral structures(p648).It is also relevant to note that evidence exists which shows a correspondence between a structure's mobility and its ability to discriminate orally presented forms(Silverman(1961) for instance).All oral areas,as we have seen,are relatively densely innervated and their receptive fields are small.The order of relative sensitivity found by Ringel in his two point discrimination experiments " approximates the progression obtained when the structures are ordered on a basis of relative size of their cortical and thalamic projection areas "(Ringel and Ewanowski(1965)), i.e. the greater the size the greater the sensitivity.

Finally,Mason(1967) studied the strategies used in oral perception tests by subjects in (a) normal (b) right unilateral mandibular block and (c) bilateral mandibular block conditions and his findings tend to fit the above pattern.Under conditions (a) and (b) subjects tended to use the anterior part of the tongue to feel the inserted form's shape.Few two area contact strategies were used (for example lingual-palatal contacts).Under condition (c) the tongue again was relied on but usually in association with other oral structures-palate or teeth etc. McDonald and Aungst(1967) used an artificial palate for some tests of oral stereognosis and found no significant change in error rate when compared with non-covered palate subjects' scores.This tends to underline the sensitivity of the tongue and its importance in making oral form perceptual judgements.

In Mason's (c) condition,the total time taken for identification was increased by , on average , four minutes.In (a) and (b) conditions subjects preferred to have the objects on the tongue midline (despite the anesthetization of the right half of the tongue) a finding which tends to correspond with Ringel's(1965) observation that the midline of all structures studied for two point discrimination was the most discriminate.Under bilateral mandibular block,all Mason's subjects used pressure manipulation on the centre and posterior tongue as the principal identification source.

3.0 It has been hypothesized (Rutherford, 1967) that a child's acquisition of phonology and ability to combine sounds into linguistically meaningful patterns is accomplished to some extent via sensorimotor associations. To acquire an adequate communication system would, therefore, according to this hypothesis, depend on the "anatomical and functional integrity of sensory receptors and their concomitant interconnections through the peripheral and Central Nervous System." (Zlatin, 1972, p.371). In the previous section it has already been indicated how a positive relationship exists between orosensory perception and speech production. People deprived of orosensory sensation either pathologically or hereditarily, provide some support for the above hypothesis and provide further data on those aspects of speech production which seem to be most affected by such deprivation.

Zlatin (1972) studied a child from the age of 11 to 47 months who had congenital sensory impairment. M (the child studied) was insensitive to touch, pain and temperature in the oral cavity due to the absence of peripheral sensory endings there. Despite this, on speech intelligibility tests at the end of the period of study, a group of 34 adult listeners could, as a group, judge correctly 84% of the child's words used in the test sentences.

The kinds of errors and difficulties which most commonly arose in the child's speech show a rather significant pattern. An analysis of feature violations at 41 months revealed that place errors accounted for 72% of articulatory errors made (manner violations accounted for 9%; voicing violations for 6% and oro-nasal violations 0%, p.387). These place errors reflected a general lack of articulatory precision and anteriorly produced sounds were most affected. Bilabial sounds were periodically absent from M's speech and when otherwise attempted, tended to be replaced by lingual-dental or lingual alveolar sounds. The tongue tip was rarely used relatively independently of the rest of the tongue. For example, alveolar sounds would tend to be produced using the tongue blade. Again it is interesting to note that the most sensitive areas of the oral cavity (and the most densely innervated) - the anterior areas of the tongue and the lips in particular - are the areas where a major proportion of articulatory errors occur. Few errors were made on vowels or velar consonants; fricatives being a major source of place and manner errors (e.g. f→d; v→b; ð→d, p.386). Errors like these in an experimentally induced state of orosensory deprivation were found by Gammon et al. (1971) where, in the labial and alveolar regions, alterations in fricatives, affricatives and stops occurred. Mason's (1967) observations, cited previously, of subjects in oral stereognosis tests using the centre and posterior tongue for form manipulation <u>only</u> when the whole of the anterior part of the tongue was anaesthetized, underlines how important this highly sensitive and mobile area is during speech production.

3.1 A series of experiments conducted by MacNeilage, Rootes and Chase (1967) provides further evidence of the importance of intact orosensory feedback mechanisms in speech. They intensively studied a 17 year old female with severe impairment of orosensory capacity and motor control. She felt no sensation of pain in the oral capacity when tested with both sharp and blunt probes (p.451). Her speech was virtually unintelligible, the impairment of her articulation being significantly more marked for consonants as opposed to vowels (in fact, her speech has been restricted "primarily to the production of vowel sounds most of her life" (p.450), this impairment being related to a "marked inability to organize movements of the lips and tongue" (451)). Neurological examination showed no direct damage to the motor system and no structural or functional deficiencies in the muscles. Tests of oral stereognosis, two-point discrimination and tactile localization, however, showed considerable deficiencies in orosensory functioning.

Her production of various sounds was phonetically analyzed. As in the previously cited study, the major source of error was in the anterior oral cavity, consonants being most severely affected, fricatives and stops especially. For example, her voiced stops were never judged as correct and voiced fricatives were seldom correct. Consonant clusters were almost never produced correctly. Cinefluorographic study revealed that the anterior part of the tongue never moved independently (p.461) which resulted in the girl not being able to articulate /e/ or affricates. This is somewhat similar to M's articulation, a further similarity being the girl's use of tongue body movements alone to produce /t/ and /n/.

3.2 One of the problems with studies of such subjects is that of separating possible motor effects on articulation from sensory based effects. Scott et al. (1971) conducted experiments to assess the effect of clinically known motor disorders on speech and compared them with speech production under experimentally induced sensory deprivation. Their general finding was that each group produced "unique kinds of articulatory errors" (p.823), in particular, "the loss of peripheral sensory information normally used in the control of speech results in a unique constellation of articulatory deviations" (p.827), so it is possible to distinguish speakers who have sensory based dysfunction from those with motor system damage.

Of special interest here are the kinds of errors made by the sensory-deprived group. Both groups had no difficulty producing phonetically acceptable vowels. Of the consonant productions, fricatives and stops suffered most and the error patterns from these two types of articulation in particular could differentiate between dysarthric and sensory-deprived speakers. Bilabials tended to be produced unilabially (using the lower lip) or with an extended release phase which produced an affricated release. Alveolar and velar stops were retracted in half the potential occurences (a tendency also noted by MacNeilage et al. (1967) in their subject). Fricatives, /s/ and /ʃ/ were less close and more retracted than in normal speech. Another recurrent finding was the inability of the sensory-deprived subjects to selectively shape the tongue apex.

3.3 It is obviously not a particularly easy task to find people who exhibit particular sensory dysfunctions so most investigations of orosensory mechanisms to date are based on experimental situations using normal subjects who are selectively deprived of some aspect of their orosensory perceptual mechanism. The patterns we have seen emerging from the previous sections again becomes evident when experimental sensory deprivation studies are considered. Sensory deprivation is usually induced using nerve block anaesthesia, this condition being the "block" condition. In the experiments to be considered here, sensation in all surface oral receptors in the supraglottal region are anaesthetized except for the pharynx and posterior one third of the tongue.

Scott and Ringel (1971) studied two subjects' production of 24 bisyllabic words and, using a classification system based on that of Peterson and Shoup (1966), collected data on "secondary articulatory parameters such as air release, laryngeal action, apex shape and lip shape as well as primary manner and place parameters". (p.807).

Again, stop manner errors tended to result from insufficient closure so that 4 out of 5 of the errors for intended initial /b/ resulted in bilabial fricative /β/. Bilabial stops were unilabially produced using the lower lip in approximately half the cases and voiceless stop release in initial position tended to be affricated. (p.808-9). Alveolar and velar stops, /t d k g/, were often produced with a retracted point of closure; alveolar stop closure tending to be effected using a comparatively large tongue surface area.

Fricatives /s,ʃ/ were almost always less closely produced with loss or diminution of fine control over tongue blade and apex manipulation in the formation of the necessary constrictions.

Vowels are generally considered to be resilient to anaesthetic - on the hypothesis (Ladefoged, 1967; Schleisser and Coleman, 1968) that they are more dependent on kinaesthetic feedback. However, Scott and Ringel did find certain changes in vowel production. Those vowels which need labial adjustment were frequently changed - /ɔ o u/ being less rounded in half the cases and vowels specified as high-front or high-back were found to be less close (p.812). Vowels in the environment of consonants which retracted under the block condition also tended to retract. A "slight" tendency for vowels to adopt a more neutral vocal tract configuration was also noted (p.813). This was of no perceptual significance though, but the hypothesis that vowels are totally dependent on only kinesthetic monitoring channels needs reconsideration.

In a further experiment, Putnam and Ringel (1972) studied the bilabials /p b m/ in initial, intervocalic and inter-consonantal position and found them unilabially produced in all cases, with incomplete closure for /p/ in all /sp/ clusters (p.534). In normal and block conditions /p b m/ could be distinguished by their lip opening rates in initial position, /p/ being fastest and /m/ slowest. /p/ changed phonemically in intervocalic position and in double and treble clusters, to /ɸ/. Lip opening rates also distinguished /p b m/ with /p/ again being swiftest and /m/ slowest. The voiceless bilabial plosive seems to be most dependent on sensory information, possibly because such information may be needed for determining intraoral air pressure which can, inter alia, distinguish /p/ from /b/ and /m/ (Malecot, 1968).

Again, vowels were altered under block conditions. After unilabial stop release, /æ/ and /i/ were less open and /u/ was not rounded (p.534). Another difference was that /æ, i/ maximum lip opening areas were half their control values and /u/'s maximum lip opening area was 2 to 3 times the control value. (p.536). This again shows the need for close examination of vowels in sensory block conditions to try and establish how much coarticulation with deviant consonants might be causing these effects and whether or not tactile deprivation significantly effects vowel production. The explanation of the relative integrity of vowel production under sensory deprivation may not be simply because such deprivation does not affect vowels to the same extent as consonants but may perhaps be because vowels can be produced with a greater degree of under- or overshoot which is not perceptually significant than consonants which, in the case of fricatives for example, require more accurate articulatory adjustments. This speculation could profitably be tested.

The tendency for fricatives /sʃ/ to be produced with a closer more retracted articulatory positioning is noted by Horii et al. (1973) who made acoustic analyses of speech produced under anaesthetic. They found significant alterations in the spectral characteristics of /sʃ/. Spectral maxima for /s/ moved down in frequency by up to 3,000 HZ and for /ʃ/ by up to 1,000 HZ (73). Relative amplitude of vowel formants were also affected though formant centre frequencies were not significantly altered.

To sum up, consonants are, as a class, more severely affected by sensory deprivation than vowels - though the effect of sensory deprivation in vowels needs reconsideration. Among consonants, fricative and affricative errors of manner of production generally outnumber such errors in plosives, laterals, glides and nasals (Gammon et al., 1971, p.277). The most frequent errors in Gammon et al.'s study of articulation under block anaesthesia were:

(a) Substitution of plosives for fricatives (in 302 out of 334 intended fricative productions).

Next were

(b) Substitution of labiodentals for bilabials (in 319 out of 330 intended bilabial productions) p.279.

Place errors generally involve a backwards (labiodental → alveolar) articulatory movement and manner errors generally involve an open to closed movement (fricatives / affricatives → plosives). The majority of place errors occur in the anterior part of the oral cavity, in particular, in the bilabial/labiodental area (Gammon et al. 1971, p.276). This suggests that more articulatory precision and, therefore, more intact orosensory feedback may be necessary for such articulations as compared with, for instance, velar or palatal productions.

4.0 Although much more investigation of orosensory feedback mechanisms is necessary before their role in speech production is comprehensively established, it is of interest, at this point, to consider the data presented here in relation to some theories of speech production which form the general conceptual framework within which a large amount of present day phonetic research is conducted.

Fairbanks (1954) posited a model of speech production which included the idea of speech as a servomechanism (cf. Chase, 1967). According to his model, speech output - in terms of auditory and orosensory feedback - was continually monitored and sampled and a high-level comparator matched it with the input of the speech mechanism. Any differences between the two would trigger an error correction signal which would modify outgoing high-level commands to the articulators (p.135). At first glance, the disruptions caused by sensory deprivation would seem to support this. However, the model would predict much more radical breakdown because of its virtually total dependence on closed loop feedback from the periphery. The fact, however, that speech remains highly intelligible under extensive orosensory feedback deprivation does not support Fairbanks' model. It could be argued that auditory feedback compensates for loss or diminution of orosensory feedback but Schleisser and Coleman (1968); Gammon et al. (1971) and Ringel and Steer (1963) found that even with both orosensory deprivation and auditory masking speech still remains highly intelligible. Coarticulation studies showing anticipatory articulator movements also suggest some autonomous high-level motor command system involved in speech production (Sussman et al., 1973; Ohman, 1966; Daniloff and Kim, 1971).

If speech is considered as a "skilled bodily action" as Ladefoged (1967) suggests (p.162) then the degree of speech efficiency which is exhibited would depend on how well the speaker could process and coordinate incoming sense data with outgoing motor activity. Ladefoged saw 3 kinds of feedback necessary for this (a) auditory (b) tactile (c) kinesthetic (p.163). This kind of "efficiency" is what orosensory tests purport to measure and the distinctions with regard to articulatory proficiency between various groups of speakers discussed in Section 3 tend to support the suggestion of speech efficiency being to some extent dependent on the integrity of this sensory monitoring and integrating system. Ladefoged also assigned different sensory monitoring systems to different classes of articulation (p.164-5). He saw auditory feedback as the principal monitoring system for preserving vowel quality, nasality and pitch and orosensory feedback as the monitoring system for consonants. Interestingly, Perkell (1969) makes a somewhat similar speculation about such a division of peripheral mechanisms. He posits 2 underlying neuromuscular systems involved in speech which deal with different kinds of peripheral feedback. Vowels which principally depend on the tongue's extrinsic muscles use acoustic and kinesthetic information principally and consonants, using extrinsic muscles and faster intrinsic muscles, depend more on tactile feedback and intraoral air pressure feedback.

Both Ladefoged and Perkell seem to tacitly assume that there is closed-loop control of speech and both also assume that if certain articulations are not affected by orosensory deprivation, it is probably because some other feedback mechanism is being relied on. Ringel (1970) believes that such attempted divisions of monitoring responsibilities are not at present justified by the evidence. (Cf. Studdert-Kennedy and Shankweiler, 1970).

According to Henke (1967) the normal temporal succession of speech movements cannot be effected without the use of peripheral feedback. For instance, he would posit that without tactile feedback to indicate bilabial closure, the command system which waits on this signal before proceeding would have its time schedule disrupted and the normal articulatory movement (in time) from one position to another would be changed. Certain evidence would seem to support this: the fricated release of initial /p/ and the correlated delay in VOT for the succeeding vowel which Putnam and Ringel found (1972). Again it is assumed that use of proprioceptive or auditory feedback compensates for loss of tactile or pressure sensitivity.

4.1 The extreme opposing view of the speech production mechanism is that speech is a completely open-loop system. As MacNeilage (1970) says, this system would "not have to wait for information associated with actually reaching the previous location in order to control the following movement appropriately" (p.189). With such an autonomous motor command system one would predict no disruption to speech under conditions of orosensory deprivation and, as we have seen, this is not the case. A point between these extremes yields a more satisfactory explanatory model.

The work of the Haskins experimenters on the motor theory of speech perception (Liberman et al., 1963, 1967) posits that we learn to associate our speech output with our perception of that output. Rootes and MacNeilage (1967) found such an association in their study of the congenitally deprived girl. Among stop consonants and front vowels "the most efficiently produced phones (/i/ and /b/) are also the ones most successfully identified". (p.317). However, her speech was "almost completely unintelligible" to the authors though her perception of the normal speech of others was apparently normal (p.317). The problem, of course, with subjects who are deprived of normal sensory input over a long period as opposed to short term deprivation, is that they lack normal coordinated sensorimotor experience and, if, as Ladefoged (1967) and Bishop et al. (1973) suggest, speech efficiency is related to a speaker's ability to integrate sensory information with motor patterns of speech, one would expect severe disruption in the case of the girl mentioned above. The whole subject of normal orosensory experience in the acquisition of speech and maintenance of its normal efficiency needs investigation. In the experimental investigations discussed earlier, short term deprivation could be compensated for by the previously established integrity of the speech mechanism as the high level of intelligibility shows. Such an investigation should also include a study of the posited space-coordinate systems in speech users' brains (MacNeilage, 1970) which tests such as oral stereognosis seem to suggest exist.

4.2 A model which includes both open and closed loops systems in speech was suggested by Liberman et al. (1967) to work in the following manner:

"The instructions might be of two types, 'on-off' or 'go-to', even in a maximally simple model. In the one case, the affected muscle would contract or not with little regard for its current state (or the position of the articulator it moves): in the other, the instruction would operate via the gamma efferent system to determine the degree of contraction (hence the final position of the articulator, whatever its initial position.)" (p.447).

It is generally accepted that there is an autonomous open loop system involved in speech, primarily to encode at the high level not just one articulatory segment at a time but a succession of such segments posited as being about 7 syllables long (Kozhevnikov and Chistovich, 1965; Gammon et al., 1973; Laver, 1968) and within this general framework, Liberman et al.'s hypothesis seems plausible and there is a certain amount of evidence to support it. According to their hypothesis, in an initial /p/ production, a simple open-loop, 'on-off' closure gesture would be necessary, but in order to go to the following articulatory position the command system must know where the relevant articulators are moving from in order to calculate the distance in time which they have to cover. Putnam and Ringel's (1972) findings would seem to corroborate this. In their study, referred to earlier, initial /p b m/ suffered no major changes in manner of production under anaesthetic whereas, for instance, intervocalic /p/ and /p/ after /s/ changed to fricatives and /r/ and /u/ following initial /p/ were often not rounded (p.539). Kozhevnikov and Chistovich (1965) showed that lip closure velocity in bilabials was directly proportional to the amount of lip opening of the preceding vowel. The anaesthetic, we could posit, disrupted the 'where from' information.

Scott and Ringel (1971) also support this position with some interesting observations. Under anaesthetic, subjects can make basic manner gestures not significantly affected by anaesthetic such as: complete closure of the vocal tract, opening and closing of the palate, basic vocal tract configurations for vowels and close vocal tract constrictions (p.815). This provides a basis for positing high-level encoding of motor commands for specific targets which would in the main be gross manner gestures. Closed loop control via tactile, pressure, proprioceptive and auditory mechanisms then becomes necessary to refine these basic gestures, particularly, as we have seen, in articulations involving relatively more precise articulator positioning (Liberman et al., (1967)). The lack or loss of refinement of apical tongue gestures noted previously in pathological or congenital sensory deprived speakers also suggests that we can make a distinction between basic gross gestures and those which are more refined.

Moll and Shriner (1967) show that palatal position varies as a function of segment type so on-off commands cannot be expected at the periphery. They posit environmental restraints as the reason for this but maintain that, at the high level, the command can be posited as binary.

A further complication involved and a further demonstration of the role of closed loop control has been pointed out by MacNeilage (1970) referring to work on speech initiation gestures by the jaw (for instance, MacNeilage, Krones and Hanson, 1970):

"although the amount of opening of the jaw varies over a range of several mms. under non-speech conditions, the jaw positions adopted for a given initial segment shows very little variation (1-2 mms.) with repetitions of the same utterance. For this to be so, the production mechanism must take into account the pre-speech position of the jaw; and make a speech-initial jaw movement contingent upon the pre-speech position." (p.192).

The problem of explaining peripheral variance in models which posit invariant high-level phonological inputs will probably involve specifying sensory-feedback systems and their modifying role at the periphery and most speech production models discussed here could probably include in their investigations of the motor control of speech as an <u>output</u> oriented device a consideration of input influences on that output (Lindblom (1972); Tatham (1969)). A fruitful approach could be an examination of the influences of orosensory deprivation on coarticulation which could throw light on the complex relationship between autonomous central motor control and peripheral sensory mechanisms. The degree to which a relatively independent articulator can compensate for overshoot or undershoot may depend to some extent on the sensitivity of the articulators involved to peripheral sensory information.

Previous sections have demonstrated the existence of methods of differentiating classes of articulations (correlated with major phonological classes) and some feature specifications. The distinction found between vowels and consonants is the obvious example of the former and the distinction between the anterior and posterior parts of the oral cavity (correlating to some extent with the anterior/coronal phonological division, (Gammon et al., 1973)) exemplify the latter.

Detailed investigation of bilabials and fricatives, for example, which seem most effected by deprivation and of various place differentiations between articulations and comparison of this data with phonological classificatory systems could be enlightening in view of the present interest in phonological natural classes. The general finding that anaesthetic may selectively effect certain classes of sounds (Horii et al. 1973) and their intelligibility gives further reason for such investigations. The theory of Markedness uses, as part of its justification for assigning m or u markers to certain phonological entities a concept of relative complexity - in, for example, the number of features needed to specify a segment or a class. Studies of orosensory mechanisms also involve the concept of complexity of different articulations (for e.g. fricatives as opposed to nasals) and also could provide a relatively concrete procedure for defining relative complexity[1] in terms of, for example, articulators involved, dependence on various feedback mechanisms etc.

The output of the phonology provides a set of instructions to the articulators on the assumption that there exists some idealized speech norm - some norm of articulatory proficiency. Studies of speakers with orosensory deprivation, whether experimental, pathological or congenital, could help establish those aspects of speech production (tactile, proprioceptive feedback, etc.) without which articulation cannot be classed as 'normal'. One could investigate the previously posited basic manner gestures with a view to this.

[1] and also for defining the amount of relative articulatory precision which may be necessary for speech perception purposes.

4.4 Much more data needs to be collected on the nature and function of sensory input mechanisms and their specific role in speech production. There is virtually no information at present on the role of sensory feedback on compensatory articulator activity. The effect which orosensory deprivation of a particular articulator has on the activity of synergistic articulators involved in producing the same articulatory gesture also needs investigation.

In order to establish the relative importance of sensory mechanisms in specific articulatory gestures (e.g. in bilabials), the other known parameters involved must be simultaneously examined. The intraoral pressure mechanism posited by Malecot (1968) has yet to be conclusively established. The relative roles in articulation of proprioceptive, tactile, pressure and auditory feedback mechanisms also need thorough investigation. Other techniques such as palatography or x-ray filming would be necessary for collecting quantitative data on such phenomena as 'observed' tongue retraction etc. and possible alterations in place of articulation generally.

One of the problems with EMG studies of muscle activity in conditions of orosensory deprivation and a general problem in all orosensory deprivation studies is an anatomical one. The difficulty in identifying individual muscles for EMG investigations using either surface-cup or needle electrodes is well known (Harris, 1970). Work by, for example, Ohman et al. (1965) aimed at providing a reliable method of ensuring proper positioning of electrodes can help here. More specifically, the role of anaesthetic in experimental deprivation experiments is a problem which needs investigation. Little is known of possible leakages of the anaesthetic to motor nerves, neighbouring those selected for anaesthetic. Usubiaga et al. (1967), demonstrated how certain commonly used anaesthetics (procaine, lignocaine) enter the cerebrospinal fluid extremely rapidly (within 20 minutes) and this may possibly effect the brain speech-controlling mechanism. However, considering our current knowledge of the subject the similarity already established between experimentally deprived subjects' speech and the speech of congenitally or pathologically deprived subjects' speech tends to suggest this influence is not absolutely crucial.

REFERENCES

Bishop, M. E., Ringel, R. L., & House, A. S. (1973) — Orosensory perception, speech production and deafness. Journal of Speech and Hearing Research, 16, 257 - 266.

Bosma, J. F. (Ed.) (1967) — Symposium on Oral Sensation and Perception. Springfield, Ill.: Charles Thomas.

Bosma, J. F. (Ed.) (1970) — Second Symposium on Oral Sensation and Perception. Springfield, Ill.: Charles Thomas.

Bosma, J. F. (Ed.) (1972) — Third Symposium on Oral Sensation and Perception. Springfield, Ill.: Charles Thomas.

Bowman, J. P. & Combs, C. M. (1969) — The cerebrocortical projection of hypoglossal afferents. Experimental Neurology, 23, 291 - 301. Cited in Konigsmark (1970).

Chase, R. A. (1967) — Abnormalities in motor control secondary to congenital sensory defects. In Bosma (Ed.), 302 - 309.

Edwards, E. P. (1970) — A comparative study of form perception, kinesthetic and spatial orientation abilities in articulatory defective and normal speaking children. MA dissertation, Purdue University. Cited in Bishop et al.

Fairbanks, G. (1954) — Systematic research in experimental phonetics: 1. A theory of the speech mechanism as a servosystem. Journal of Speech and Hearing Disorders, 19, 133 - 139.

Gammon, S. A., Smith, P.J., Daniloff, R. G. & Kim, C. W. (1971) — Articulation and stress/juncture production under oral anaesthetisation and masking. Journal of Speech and Hearing Research, 14, 271 - 282.

Hardy, J.C. (1970) Development of Neuromuscular Systems Underlying Speech Production. ASHA Reports, 5,49 - 67.

Harris, K.S. (1970) Physiological Measures of Speech Movements: EMG and Fiberoptic Studies. ASHA Reports 5, 271-282.

Henke, W. (1967) Preliminaries to Speech Synthesis Based Upon An Articulatory Model. Conf. Speech Compression Processing, Bedford, Mass.: Office of Aerospace Research, ACRL, USAF.

Henneman, E. (1968a) Cerebellum, In Mountcastle (1968), 1771 - 1789. Cited in Hardy (1970)

Henneman, E. (1968b) Organisation of the Motor Systems - A Preview. In Mountcastle (1968), 1675 - 1680. Cited in Hardy (1970).

Horii, Y., House, A.S., Kung-Pu, Li & Ringel, R.L. (1973) Acoustic Characteristics of Speech Produced Without Oral Sensation. Journal of Speech and Hearing Research, 16, 67 - 77.

Jerge, C.R. (1967) The Neural Substratum of Oral Sensation. In Bosma (1967), 63 - 83.

Kaplan, H.M. (1960) Anatomy and Physiology of Speech. McGraw-Hill, N.Y.

Kawamura, Y. (1970) A Role of Oral Afferents For Mandibular and Lingual Movements in Bosma (1970) 170 - 192.

Konigsmark, B.W. (1970) Neuroanatomy of Speech. ASHA Reports,5, 3 - 19.

Kozevnikov, V.A. Speech: Articulation and Perception. Joint Publications Research Service, U.S. Dept. of Commerce, Washington, D.C.

Ladefoged, P. (1967) Three Areas of Experimental Phonetics. Oxford University Press.

Laver, J. (1968) Phonetics and the Brain. Work in Progress 2, Edinburgh University, 63 - 75.

Liberman, A.M., Cooper, F.S., Harris, K.S., & MacNeilage, P.F. (1963) A Motor Theory of Speech Perception. Proc. Speech Comm. Seminar, R.I.T. Stockholm, D3.

Liberman, A.M., Cooper, F.S., Shankweiler, D.P. & Studdert-Kennedy, M. (1967) Perception of the Speech Code. Psychological Review 74, 431 - 461.

Laver, J. (1968) Phonetics and the Description of Language in _Phonetic Sciences and Linguistic Theories_, Proc. VIIth. Int. Congr. Phonetics Sciences, Montreal (1971). Mouton, The Hague. 63 - 97.

Locke, J.L. (1968) Oral Perception and Articulation Learning. _Percept. and Motor Skills_, 26, 1259 - 1264.

MacNeilage, P.F. (1970) Motor Control of the Serial Ordering of Speech. _Psychol. Review_, 77, 3, 182 - 196.

MacNeilage, P.F., Krones, R., & Hanson, R. (1970) Closed Loop Control of the Initiation of Jaw Movement for Speech. _Journal of the Acoustical Society of America_, 47, 104.

MacNeilage, P.F., Rootes, T.P., & Chase, R.A. (1967) Speech Production and Perception in a Patient with Severe Impairment of Somesthetic Perception and Motor Control. _Journal of Speech and Hearing Research_, 10, 449 - 467.

Malecot, A. (1968) The Force of Articulation of American Stops and Fricatives as a Function of Position. _Phonetica_, 18, 95-102.

Mason, R.M. (1967) Studies of Oral Perception Involving Subjects with Alterations in Anatomy & Physiology. In Bosma (1967), 294 - 301.

McDonald, E.T. & Aungst, L.F. (1967) Studies in Oral Sensorimotor Function. In Bosma (1967), 202 - 220.

McDonald, E.T. & Aungst, L.F. (1970) An Abbreviated Test of Oral Stereognosis. In Bosma (1970), 384 - 390.

Moll, K.L. & Shriner, T. (1967) Preliminary Investigation of a New Concept of Velar Activity During Speech. Cleft Palate J.4.

Mountcastle, V.B. (Ed.) (1968) Medical Physiology Vol.11, 12th Edition: Mosby, St. Louis.

Mountcastle, V.B. & Henneman, E. (1952) The Representation of Tactile Sensibility in the Thalamus of the Monkey. J. Comp. Neurol, 97, 409 - 439: cited in Konigsmark (1970).

Ohman, S.E.G. (1966) Co-articulation in CVC Utterances: Spectrographic Measurements Journal of the Acoustical Society of America, 39, 151 - 168.

Ohman, S.E.G., Leanderson, R., & Persson, A. (1965) Electromyographic Studies of Facial Muscles During Speech. STL - QPSR, 3, 1 - 11.

Perkell, J.S. (1969) Physiology of Speech Production: Results and Implications of a Quantitative Cineradiographic Study. Cambridge, Mass.: M.I.T. Press.

Peterson, G.E. & Shoup, J. (1966) A Physiological Theory of Phonetics. Journal of Speech and Hearing Research, 9, 5 - 67.

Putman, A.H.B. & Ringel, R.L. (1972) Some Observations of Articulation During Labial Sensory Deprivation. Journal of Speech and Hearing Research, 15, 529 - 542.

Ringel, R.L. (1970) Oral Sensation and Perception: A Selective Review. ASHA Reports, 5, 188 - 205.

Ringel, R.L., Burk, K.W. & Scott, C.M. (1968) Tactile Perception: Form Discrimination in the Mouth. Brit. J. Comm. Dis., 3, 410

Ringel, R.L. & Ewanowski, S.J. (1965) Oral Perception: 1. Two Point Discrimination. Journal of Speech and Hearing Research, 8, 389 - 398.

Ringel, R.L. & Fletcher, H.M., (1967) Oral Perception: 3. Texture Discrimination. Journal of Speech and Hearing Research, 10, 642 - 649.

Ringel, R.L., House, A.S., Burk, K.W., Dolinsky, J.P., & Scott, C.M. (1970) Some Relations Between Orosensory Discrimination and Articulatory Aspects of Speech Production. Journal of Speech and Hearing Disorders, 35, 3 - 11.

Ringel, R.L., Saxman, J.H., & Brooks, A.R. (1967) Oral Perception: 2. Mandibular Kinesthesia. Journal of Speech and Hearing Research, 10, 637 - 641.

Ringel, R.L. & Steer, M.D. (1963) Some Effects of Tactile and Auditory Alterations on Speech Output. Journal of Speech and Hearing Research, 6, 369 - 378.

Rootes, T.P. & MacNeilage, P.F. (1967) Some Speech Production and Perception Tests of a Patient With Impairment in Somesthetic Perception and Motor Function. In Bosma (1967). 310 - 317.

Rosenbek, J.C. & Wertz, R.T. (1973) Oral Sensation and Perception in Apraxia of Speech and Aphasia. Journal of Speech and Hearing Research 16, 22 - 36.

Rutherford, D. (1967) Auditory-Motor Learning and the Acquisition of Speech. Am. J. Phys. Med., 46, 245 - 251. Cited in Zlatin (1972).

Schliesser, H.F. & Coleman, R.O. (1968) Effectiveness of Certain Procedures for Alteration of Auditory and Oral Tactile Sensation for Speech. Percept. and Motor Skills, 26, 275 - 281.

Scott, C.M. & Ringel, R.L. (1971a) — Articulation without Oral Sensory Control. Journal of Speech and Hearing Research, 14, 804 - 818.

Scott, C.M. & Ringel, R.L. (1971b) — The Effects of Motor and Sensory Disruptions on Speech: A Description of Articulation. Journal of Speech and Hearing Research, 14, 819 - 828.

Shelton, R.L. & Hetherington, J.J. (1967) — Testing Oral Stereognosis. C 12 in Bosma (1967), 221.

Silverman, S. (1961) — Oral Physiology. Mosby, St. Louis.

Studdert-Kennedy, M. & Shankweiler, D. (1970) — Hemispheric Specialization for Speech Perception. Journal of the Acoustical Society of America, 48, 579 - 594.

Sussman, H.M., MacNeilage, P.F. & Hanson, R.J. (1973) — Labial and Mandibular Dynamics During the Production of Bilabial Consonants: Preliminary Observations. Journal of Speech and Hearing Research, 16, 397 - 420.

Tatham, M.A.A. (1969) — On the relationship between Experimental Phonetics and Phonology. UELCOP, 5, 20 - 26.

Usubiaga, J.E., Moya, F., Wikinski, R. & Usubiaga, L.E., (1967) — Relationship Between the Passage Of Local Anaesthetics Across the Blood-Brain Barrier and Their Effects on the Central Nervous System. Brit. J. Anaesthesia, 39, 943 - 947.

Zlatin, M.A. (1972) — Development of Speech, Language, Auditory and Oral Function in The Presence of Congenital Sensory Neuropathy. In Bosma (1972), 381 - 390.