Some formal properties of phonological redundancy rules.

Stephan Braun*

1. Introduction.

Redundancy is a well-known phenomenon of phonemes or phonological matrices within the framework of the distinctive-feature theory of JAKOBSON and HALLE [1]. Redundancy in this theory means that the specification (either + or -) of certain features of a phoneme is predictable given the specifications of certain other features of the same phoneme and/or of neighbouring phonemes of a phoneme sequence. These restrictions on feature specifications are usually expressed by "redundancy rules". E.g. in English all nasal phonemes are voiced which is expressed by a rule [+nasal] → [+voiced], to be read as "each phoneme which is specified [+nasal] must also be specified [+voiced]". Among the redundancy rules usually two main types are distinguished. Those like the one just mentioned which express a restriction valid for each phoneme of a language, independent of possible neighbouring phonemes, will be called "phoneme-structure rules" (P-rules) in this paper. Besides them, there are rules expressing restrictions on the admissible phoneme sequences of the language, e.g. in English no [+consonantal] segment can follow a morpheme-initial nasal; they will be called (as usual) "morpheme-structure rules" (M-rules). In the paper of STANLEY [2] the former are called segment structure rules and the latter sequence structure rules.

The aim of the present paper is to investigate the properties of phonological redundancy rules on a mathe-

* 8 München 2
Arcisstrasse 21
Technische Hochschule
West-Germany

matical basis. Some of the problems arising in connection with redundancy rules in phonology have been mentioned already in the work of HALLE [3] where they are treated essentially on a linguistically intuitive basis. The paper of UNGEHEUER [4] on the mathematical properties of the distinctive feature system (using Boolean algebra by virtue of the fact that every feature can have exactly two specifications) mentions redundancy without going, however, into details. A very thorough and comprehensive treatment of the subject is given in the already mentioned paper of STANLEY where a formal way of arguing is used though no mathematical proofs are given. At any rate, STANLEY's results show that a formalized treatment of phonological redundancy is sensible. Most recently, redundancy rules have been discussed in the work of CHOMSKY and HALLE [5] .

The results of the present paper essentially confirm - as far as the questions are the same - the results of STANLEY being, however, somewhat more precise than his. The main result is that the complete set of P-rules for a set of fully specified phonemes can be derived from the prime implicants of a certain Boolean function and thus computed without recurrence to linguistic intuition, given only the set of phonemes. Algorithms for this task can be found in the mathematical literature (e.g. McCLUSKEY [6]). This formulation then also allows in a simple way to test intuitively found P-rules for compatibility with a given set of phonemes. No hierarchy of the features need be assumed for this. Moreover, it is shown that phoneme sequences can be treated formally like single phonemes (with a higher number of features); thus all results for single phonemes hold for phoneme sequences as well, and M-rules are not essentially different from P-rules. Furthermore, some ideas are given how to compute from a set of P-rules another set of rules which generate just the non-redundantly specified matrices, i.e. the lexicon; these rules are called "lexicon rules" (L-rules). Finally, two questions connected with the introduction into phonological matrices of blanks for redundant specifications are discussed, viz. "When do different matrices remain distinct - in the technical sense of [2] , p.408 - after introduction of blanks?" and "Is the position of blanks in matrices uniquely determined by the redundancy rules alone or has an order of application of the rules to be taken into account?". It is shown that both distinctness and uniqueness are guaranteed if a hierarchy (a total ordering) is introduced among the features and if the feature on the right hand side of a rule is required to have higher rank with respect to this hierarchy - e.g. usually [voiced] is given higher rank than [vocalic] - than any feature on the left hand side of the rule. Counterexamples show that neither distinctness nor uniqueness necessarily hold if this requirement is not met.

Phoneme-structure rules are discussed in Sec. 2, morpheme-structure rules in Sec.3, lexicon rules in Sec.4 and matrices with blanks in Sec.5.

2. Phoneme-structure rules.

As mentioned in the Introduction a phoneme-structure rule (P-rule) is a statement predicting certain feature specifications of a single phoneme given other feature specifications of this phoneme. In order to formalize this concept some notational conventions will be introduced. Let $\mathbb{B} = \{B_1, \ldots, B_p\}$ a set of p fully specified phonemes and $\mathbb{F} = \{f_1, \ldots, f_n\}$ the set of n features and $\mathbb{A} = \{+, -\}$ the set of the two possible specifications. Any phoneme $B \in \mathbb{B}$ can then be written as a set of n ordered pairs: $B = \{\beta_1 f_1, \ldots, \beta_n f_n\}$ with $\beta_i \in \mathbb{A}$ for $i=1,\ldots,n$. Every set of $m \leq n$ ordered pairs $\beta_i f_i$ containing each feature only once will be called "phonemic set"; the phonemes of $\mathbb{B}$ are thus special phonemic sets. This set-theoretic notation for phonemes is almost identical to the usual linguistic notation and will be mainly used throughout this paper; the only difference is that no ordering of the features is considered so far. It turns out that ordering of the features need be introduced only much later; for the time being it would only unnecessarily complicate the proofs.

Another notation for phonemes stems from the observation that there are exactly two specifications for each feature. The features can, therefore, be conceived of as Boolean variables taking the values true and false and a phoneme B can be written as a conjunction of these variables. E.g. $B = \{+f_1, -f_2, +f_3\}$ in set-theoretic notation is replaced by the conjunction $\hat{B}(f_1, f_2, f_3) = f_1 \wedge \bar{f}_2 \wedge f_3$ ($\bar{f}$ is the complement of f taking the value true if f takes the value false and vice versa) which takes the value true if and only if f_1 takes the value *true*, f_2 takes the value *false* and f_3 takes the value *true*. Thus *true* corresponds to the specification +, *false* to the specification − and $\hat{B}$ is formed from B by writing f_i instead of $+f_i$ and $\bar{f}_i$ instead of $-f_i$. This correspondence of $\hat{B}$ and B evidently is biunique. The whole set of phonemes is in this notation described by the Boolean function

$$(1) \qquad g(f_1,\dots,f_n) = \bigvee_{B \in \mathbb{B}} \hat{B}(f_1,\dots,f_n)$$

($\bigvee$ denotes disjunction − the logical *or*) which takes the value *true* if and only if at least one of the $B(f_1,\dots,f_n)$ takes the value *true*, i.e. if the B corresponding to $\hat{B}$ is a phoneme of $\mathbb{B}$. For the following the complement function $\bar{g}$ of g given by

$$(2) \qquad g(f_1,\dots,f_n) = \bigvee_{C \notin \mathbb{B}} \hat{C}(f_1,\dots,f_n)$$

will be of some importance. g describes the set of those phonemic sets with n features which are *not* phonemes of $\mathbb{B}$. This set which will be denoted by $\bar{\mathbb{B}}$ is in practice much larger than the set $\mathbb{B}$ since there are 2^n phonemic sets with n features while the number p of phonemes of a natural language is much smaller than 2^n for usual values of n (e.g. n=12).

A prediction for a feature specification of a single phoneme (a P-rule) is, in the set-theoretic notation, a statement of the form

$$(3) \qquad \{\alpha_1 r_1,\dots,\alpha_k r_k\} \longrightarrow \alpha r$$

with $r_i \in \mathbb{F}$, $r \in \mathbb{F}$, $r \neq r_i$ for $i=1,\dots,k$, $\alpha_i \in \mathbb{A}$, $\alpha \in \mathbb{A}$, $0 \leq k \leq n-1$, which is to be read as "if the phonemic set $a = \{\alpha_1 r_1,\dots,\alpha_k r_k\}$ on the left hand side of (3) is a part (a subset) of some phoneme B of $\mathbb{B}$ then the feature r is in B necessarily specified as α" *). Note that the condition $a \subset B$ corresponds to STANLEY's "submatrix interpretation of rule application" (cp. [2], p.413).

Now, in order for (3) to be called a prediction in a sensible sense of this word two obvious requirements must be fulfilled:

(i) a must occur in at least one phoneme of $\mathbb{B}$

(ii) α must be uniquely determined by a and r

For simplicity we add a further requirement

(iii) a must be minimal, i.e. there is no phonemic set $b \subset a$ such that b and r already suffice to uniquely determine the specification of r in B.

Since by (ii) a uniquely predicts α as specification of r there is no phoneme $P \in \mathbb{B}$ such that the phonemic set $h = a \cup \{\bar{\alpha} r\}$ (i.e. a plus the feature r specified as $\bar{\alpha}$, written as set-theoretic union) **) is a subset of P. Any phonemic set with n features containing h is, therefore, an element of $\bar{\mathbb{B}}$. A phonemic set h with this property is called <u>implicant</u> of $\bar{\mathbb{B}}$. More specifically, we define the notion of <u>prime implicant</u> of $\bar{\mathbb{B}}$:

<u>Definition 1</u>

A phonemic set $h = \{\alpha_1 r_1, \ldots, \alpha_m r_m\}$ $(1 \leq m \leq n)$ is called <u>prime implicant</u> of $\bar{\mathbb{B}}$ if and only if

(a) there is no $B \in \mathbb{B}$ such that $h \subset B$.

---

*) The case k=0 means "r is specified as α in each phoneme of $\mathbb{B}$".

**) $\bar{\alpha} = +$ for $\alpha = -$ and $\bar{\alpha} = -$ for $\alpha = +$.

(b) for every proper subset $b \subset h$ there exists a $B \in \mathbb{B}$ such that $b \subset B$.

Condition (b) of Def.1 expresses a minimality requirement on h which will turn out to be closely related to requirement (iii) above.

The name "prime implicant" for h was chosen because in the Boolean notation of eqs.(1) and (2) the conjunction h corresponding to h is a prime implicant (in the technical sense of the theory of Boolean functions) of the function $\overline{g}$, eq.(2): An implicant of a Boolean function v of n variables is a conjunction q of $m \leq n$ of these variables such that v is *true* whenever q is *true*; equivalently, if t is any conjunction of the n variables which contains q then t = *true* implies v = *true*. q is a prime implicant of v if it is an implicant of v and if every proper part s of q is not an implicant of v; equivalently, if there is at least one conjunction w of the n variables containing s such that w = *true* implies v = *false* (or $\overline{v}$ = *true*). By condition (a) of Def.1 $P \in \overline{\mathbb{B}}$ for every phonemic set P with n features with $h \subseteq P$; in Boolean notation $\hat{P}$ is any n_place conjunction containing $\hat{h}$ and $P \in \overline{\mathbb{B}}$ means $\hat{P}$ = *true* implies $\overline{g}$ = *true*. Thus $\hat{h}$ is an implicant of $\overline{g}$. Condition (b) of Def.1 in Boolean notation reads "if $\hat{b} \subset \hat{h}$ then there is a $\hat{B}$ with $\hat{b} \subset \hat{B}$ such that $\hat{B}$ = *true* implies g = *true* (or $\overline{g}$ = *false*)". Thus $\hat{h}$ is a prime implicant of $\overline{g}$.

The remarks following conditions (i) through (iii) together with Def.1 suggest a connection between prime implicants of $\overline{\mathbb{B}}$ and P-rules. This is expressed by

Theorem 1

1. From each prime implicant $h = \{\alpha_1 r_1, \ldots, \alpha_m r_m\}$ of $\overline{\mathbb{B}}$

m P-rules

$$P_j = a_j \rightarrow \bar{\alpha}_j r_j \qquad (j=1,\dots,m)$$

with $a_j = h \setminus \{\alpha_j r_j\}$ (i.e. a_j is formed from h by omitting $\alpha_j r_j$) can be derived **which** comply with conditions (i) through (iii).

2. If $P = a \rightarrow \alpha r$ ($a = \{\alpha_1 r_1, \dots, \alpha_k r_k\}$, $k \geq 0$) is a P-rule complying with (i) through (iii) then

$$h = a \cup \{\bar{\alpha} r\} = \{\alpha_1 r_1, \dots, \alpha_k r_k, \bar{\alpha} r\}$$

is a prime implicant of $\bar{\mathbb{B}}$. P is derived from h by 1., and h is uniquely determined by P.

Proof:

1. P_j evidently has the form of eq.(3). Since h is a prime implicant of $\bar{\mathbb{B}}$ and $a_j \subset h$ there is, by Def.1(b), a $B \in \mathbb{B}$ such that $a_j \subset B$. Thus, P_j complies with (i). The feature r_j omitted in a_j is in B necessarily specified as $\bar{\alpha}_j$ since it must be specified somehow and cannot be specified as α_j because then $h \subset B$ contrary to Def.1(a). Thus a_j and r_j uniquely determine α_j and (ii) is met. Suppose there is a $b \subset a_j$ such that b and r_j already uniquely determine $\bar{\alpha}_j$. Then there is, by Def.1(a), no $B \in \mathbb{B}$ containing $c = b \cup \{\alpha_j r_j\}$. But this contradicts Def.1(b) since c is a proper part of h. Thus there is no such b and P complies with (iii), too.
2. There is no $B \in \mathbb{B}$ such that $h \subset B$. For, otherwise, r is specified as $\bar{\alpha}$ instead of α in some phoneme of $\mathbb{B}$ containing a which contradicts (ii). Thus h is, by Def.1(a), an implicant of $\bar{\mathbb{B}}$. Each proper subset of h is part of a $B \in \mathbb{B}$: By (i) and (ii)

there exists a $B \in \mathbb{B}$ such that $c = a \cup \{\alpha r\}$ is a part of B. Each proper subset of c is, therefore, also a part of this B. Each proper subset of h which does not contain $\bar{\alpha}r$ is a subset of a, thus a proper subset of c, thus a part of B. Let $d = b \cup \{\bar{\alpha}r\}$ with $b \subset a$ be a proper subset of h containing $\bar{\alpha}r$. Suppose there is no $B \in \mathbb{B}$ such that d is a part of B. Then r is never specified as $\bar{\alpha}$ in all those phonemes of $\mathbb{B}$ which contain b (since $b \subset a$ and (i) there are such phonemes) but always as α. Thus $b \subset a$ and r suffice to uniquely determine α which contradicts (iii) for P. Therefore, also d is a part of some $B \in \mathbb{B}$. Thus h is, by Def.1(b), a prime implicant of $\bar{\mathbb{B}}$ and, by 1., P is derived from h.
Let $h' = \{\gamma_1 s_1, \ldots, \gamma_\ell s_\ell\}$ a prime implicant of $\bar{\mathbb{B}}$. Every P-rule derived from h' has the form $P' = a'_j \rightarrow \bar{\gamma}_j s_j$. For P to be one of these P' a comparison shows that necessarily $a'_j = a$, $\bar{\gamma}_j = \alpha$ and $s_j = r$. Then $h' = a'_j \cup \{\gamma_j s_j\} = a \cup \{\bar{\alpha} r\} = h$; thus h is uniquely determined by P.

According to Theorem 1 every P-rule for $\mathbb{B}$ complying with requirements (i) through (iii) - it seems rather obvious that a P-rule should meet these requirements - is derived from a corresponding prime implicant of $\bar{\mathbb{B}}$. The task of finding all the P-rules for $\mathbb{B}$ is, therefore, equivalent to the task of finding all the prime implicants for $\bar{\mathbb{B}}$ or, equivalently, the prime implicants of the Boolean function $\bar{g}$. This is a well-known mathematical problem which can be more or less efficiently solved on a computer using e.g. the

McCLUSKEY algorithm [6] . (The efficiency of this algorithm depends rather strongly on the number n of features; n must not be too large). Moreover, this result means that, given only the set $\mathbb{B}$ of fully specified phonemes, the discovery of P-rules for this set need not depend on linguistic intuition; the complete set of P-rules can be computed via the prime implicants of $\bar{g}$ which is, in turn, directly determined by $\mathbb{B}$.

By their connection to the prime implicants of $\bar{\bar{\mathbb{B}}}$ the P-rules are divided into equivalence classes: two P-rules will be called equivalent if and only if they are derived from the same prime implicant of $\bar{\bar{\mathbb{B}}}$. By Theorem 1.2 the connection between P-rule and corresponding prime implicant is extremely simple; thus equivalence of P-rules is easily tested by comparing the prime implicants. Moreover, the compatibility of an intuitively found P-rule with a given set of phonemes can also easily be tested: if $a \rightarrow \alpha r$ is the P-rule then $a \cup \{\bar{\alpha} r\}$ must be a prime implicant of $\bar{\bar{\mathbb{B}}}$; in particular, no phoneme of the set may contain $a \cup \{\bar{\alpha} r\}$.

Conditions (i) through (iii) for P-rules or, equivalently, the requirement that P-rules are to be derived from prime implicants of $\bar{\bar{\mathbb{B}}}$ are essentially identical to the "true generalization condition" of STANLEY ([2], p.421). In our set-theoretic notation this condition for a rule $a \rightarrow \alpha r$ reads

$$\{\alpha r\} \not\subset B \succ a \not\subset B \quad \text{for every} \quad B \in \mathbb{B}$$

($\succ$ means logical implication). By the rules of Boolean algebra this is equivalent to

$$\neg (a \subset B \wedge \{\bar{\alpha} r\} \subset B) \quad \text{for every} \quad B \in \mathbb{B}$$

($\neg$ means negation, $\wedge$ means conjunction), i.e. there is no B such that $h = a \cup \{\bar{\alpha} r\} \subset B$ which by Def.1(a) means that h is implicant of $\overline{\mathfrak{B}}$. Note that the true generalization condition is thus not equivalent to h being a prime implicant of $\overline{\mathfrak{B}}$; it does, in other words, not meet the minimality condition (iii). Because this condition has turned out in the proof of Theorem 1 to be rather convenient it is proposed that (iii) is added to the true generalization condition.

As an example consider the five labial consonants /p/,/b/,/m/,/f/,/v/ of English as given in HALLE [7], see tab.1. For simplicity only the four features

	p	b	m	f	v
strid	−	−	−	+	+
nas	−	−	+	−	−
cont	−	−	−	+	+
voiced	−	+	+	−	+

tab.1

[strident] , [nasal] , [continuant] and [voiced] are considered and the specifications [-vocalic] , [+consonantal] , [+grave] and [+diffuse] common to the five consonants are omitted. For this small example the prime implicants of $\overline{\mathfrak{B}}$ can be computed directly by means of Def.1: Assuming for convenience a fixed order of the features (e.g. that of tab.1) one has ordered sequences of the specifications + and - instead of the sets used so far. Then for each k in $1 \leq k \leq n=4$ all $\binom{n}{k} \cdot 2^k$ possible specification sequences of length k are formed and matched with tab.1. If such a specification sequence does not occur in tab.1 it is an implicant of $\overline{\mathfrak{B}}$, and it is a prime implicant of $\overline{\mathfrak{B}}$ if it does not contain any shorter implicant already found. Thus one gets five prime implicants of $\overline{\mathfrak{B}}$

(4) $\{$+strid,+nas$\}$, $\{$-strid,+cont$\}$, $\{$+strid,-cont$\}$,
$\{$+nas,+cont$\}$, $\{$+nas,-voiced$\}$

and from them ten P-rules (two for each of the prime implicants)

	+strid	$\longrightarrow$	-nas	+nas	$\longrightarrow$	-strid
	-strid	$\longrightarrow$	-cont	+cont	$\longrightarrow$	+strid
(5)	+strid	$\longrightarrow$	+cont	-cont	$\longrightarrow$	-strid
	+nas	$\longrightarrow$	-cont	+cont	$\longrightarrow$	-nas
	+nas	$\longrightarrow$	+voiced	-voiced	$\longrightarrow$	-nas

This is the complete set of P-rules for tab.1; any other redundancy rule is not a P-rule for this set.

3. Morpheme-structure rules

According to the Introduction morpheme-structure rules (M-rules) are predictions for feature specifications of single phonemes within phoneme sequences. The only difference between P-rules and M-rules then is that M-rules may (but not must) contain features of more than one phoneme of the sequence (see the example in the Introduction). It will be shown that M-rules can like P-rules be derived from the prime implicants of a suitable Boolean function. This is done by formally reducing the case of phoneme sequences to the case of single phonemes.

For sake of simplicity at first only sequences consisting of two phonemes of $\mathbf{B}$ are considered, i.e. sequences $B_1 B_2$ with $B_1 \in \mathbf{B}$ and $B_2 \in \mathbf{B}$. The n features of B_1 are denoted by $f_1, \ldots, f_n$ as before and the n features of B_2 as $f'_1, \ldots, f'_n$. Of course, f_i and f'_i denote the same phonological feature; they are distinguished merely formally to indicate their position in the phonemes of the sequence. For formal purposes, however, f_i and f'_i

may be considered to be different features, and thus we have two sets $\mathbb{F} = \{f_1, \ldots, f_n\}$ and $\mathbb{F}' = \{f_1', \ldots, f_n'\}$ of features. Uniting $\mathbb{F}$ and $\mathbb{F}'$ to form the set $\mathbb{F}^{(2)} = \mathbb{F} \cup \mathbb{F}'$ we can conceive of the sequence $B_1 B_2$ as a "phoneme of 2^{nd} degree" $B^{(2)} = B_1 \cup B_2'$ with the 2n features of $\mathbb{F}^{(2)}$ where B_2' is formed from B_2 by replacing in it f_i by f_i'. E.g. if $B_1 = \{+f_1, +f_2\}$ and $B_2 = \{-f_1, +f_2\}$ then $B^{(2)} = \{+f_1, +f_2, -f_1', +f_2'\}$. Let $\mathbb{B}^{(2)}$ be the set of all phonemes of 2^{nd} degree (i.e. the set of all admissible phoneme sequences of length two) then $\mathbb{B}^{(2)}$ is a subset of $\mathbb{B} \times \mathbb{B}'$ ($\times$ denotes the set-theoretic product) with $\mathbb{B}$ the original set of phonemes and $\mathbb{B}'$ identical to $\mathbb{B}$ except that f_i is replaced everywhere by f_i'. If every sequence of two phonemes is admissible (this probably is an only theoretical limiting case) then $\mathbb{B}^{(2)} = \mathbb{B} \times \mathbb{B}'$.

After this formal reduction of phoneme sequences to phonemes of higher degree it appears natural to assume that the M-rules will be nothing but the P-rules for the higher-order phoneme set, i.e. they are derivable from the prime implicants of $\overline{\mathbb{B}}^{(2)}$. This assumption is supported by the following:

A natural requirement for M-rules is that they reflect the restrictions on possible phoneme sequences of a language. In other words, if every sequence of phonemes is admissible then the M-rules should coincide with the P-rules for the set $\mathbb{B}$. The following theorem shows that this is indeed the case:

Theorem 2

Let $\mathbb{B}^{(2)} = \mathbb{B} \times \mathbb{B}'$. Then any implicant of $\overline{\mathbb{B}}^{(2)}$ which contains features of both $\mathbb{F}$ and $\mathbb{F}'$ is not a prime implicant of $\overline{\mathbb{B}}^{(2)}$.

Proof:

Let $T^{(2)} = T \cup T'$ with $T = \{\alpha_1 r_1, \ldots, \alpha_k r_k\}$ and $T' = \{\beta_1 s_1', \ldots, \beta_m s_m'\}$ a phonemic set containing the features $r_i \in \mathbb{F}$ $(i=1,\ldots,k)$ and $s_j' \in \mathbb{F}'$ $(j=1,\ldots,m)$; $\alpha_i \in \mathbb{A}$, $\beta_j \in \mathbb{A}$. T and T' contain features only from $\mathbb{F}$ and from $\mathbb{F}'$, respectively. If neither T nor T' is an implicant of $\bar{\mathbb{B}}^{(2)}$ then there are elements $B^{(2)}$ and $C^{(2)}$ of $\mathbb{B}^{(2)}$ such that T is a subset of $B^{(2)}$ and T' is a subset of $C^{(2)}$. $B^{(2)}$ and $C^{(2)}$ can be written as $B^{(2)} = B \cup B'$, $C^{(2)} = C \cup C'$ with B and C from $\mathbb{B}$, B' and C' from $\mathbb{B}'$. Since the features of T are all from $\mathbb{F}$ and the features of T' are all from $\mathbb{F}'$ T is a subset not only of $B^{(2)}$ but even of B; likewise, T' is a subset of C'. Therefore, $T^{(2)} = T \cup T'$ is a subset of the set $D^{(2)} = B \cup C'$. Since $B \in \mathbb{B}$ and $C' \in \mathbb{B}'$ we have $D^{(2)} \in \mathbb{B} \times \mathbb{B}'$, i.e. $D^{(2)} \in \mathbb{B}^{(2)}$ by assumption, and thus $T^{(2)}$ is not an implicant of $\bar{\mathbb{B}}^{(2)}$. That is, if $T^{(2)}$ is an implicant of $\bar{\mathbb{B}}^{(2)}$ then necessarily one of its proper subsets T and T' is an implicant of $\bar{\mathbb{B}}^{(2)}$ which shows that $T^{(2)}$ is not a prime implicant of $\bar{\mathbb{B}}^{(2)}$.

Thus, if every sequence of phonemes is admissible then the prime implicants of $\bar{\mathbb{B}}^{(2)}$ contain features only from $\mathbb{F}$ or only from $\mathbb{F}'$, i.e. they are prime implicants of $\bar{\mathbb{B}}$; because any prime implicant of $\bar{\mathbb{B}}$ evidently is a prime implicant of $\bar{\mathbb{B}}^{(2)}$ the sets of prime implicants of $\bar{\mathbb{B}}$ and $\bar{\mathbb{B}}^{(2)}$ are identical which means that the M-rules for $\mathbb{B} \times \mathbb{B}'$ coincide with the P-rules of $\mathbb{B}$. Prime implicants of $\bar{\mathbb{B}}^{(2)}$ other than those of $\bar{\mathbb{B}}$, in particular such with features from both phonemes of a two-phoneme sequence, consequently occur only if $\mathbb{B}^{(2)}$ is a proper subset of $\mathbb{B}\times\mathbb{B}'$, i.e. if not every sequence of phonemes is admissible. Thus the concept of M-rules as P-rules of a set of phonemes of higher degree is sensible, and M-rules are to be derived from the prime implicants of $\overline{\mathbb{B}}^{(2)}$ in exactly the same manner as P-rules are derived from the prime implicants of $\overline{\mathbb{B}}$.

Without proof we mention two special cases: if on the first or second position of the sequence the whole set $\mathbb{B}$ (or $\mathbb{B}'$, resp.) is possible then all the prime implicants of $\overline{\mathbb{B}}$ occur among those of $\overline{\mathbb{B}}^{(2)}$; if on the first or second position only a single phoneme is possible then M-rules with more than one feature contain features only from $\mathbb{F}'$ or $\mathbb{F}$, respectively.

As an example for M-rules assume that from the five phonemes of tab.1 the three sequences /pf/,/bv/ and /mb/ can be formed. Denote the features by strid 1,...,voiced 1, strid 2,...,voiced 2 for the first and second phoneme of the sequence, resp. In this set of phoneme pairs the specifications of the features strid 1, cont 1 and nas 2 are all fixed as -, and the remaining M-rules are

$$\begin{array}{ll} & +\text{nas } 1 \longrightarrow \{+\text{voiced } 1, -\text{strid } 2, -\text{cont } 2, +\text{voiced } 2\} \\ & -\text{nas } 1 \longrightarrow \{+\text{strid } 2, +\text{cont } 2\} \\ & +\text{voiced } 1 \longrightarrow +\text{voiced } 2 \\ (6) & -\text{voiced } 1 \longrightarrow \{+\text{strid } 2, +\text{cont } 2, -\text{voiced } 2\} \\ & +\text{strid } 2 \longrightarrow +\text{cont } 2 \\ & -\text{strid } 2 \longrightarrow \{-\text{cont } 2, +\text{voiced } 2\} \\ & -\text{cont } 2 \longrightarrow +\text{voiced } 2 \end{array}$$

where for each of the prime implicants only one rule has been given and for rules with the same left hand side the right hand sides have been combined for abbreviation.

From the preceding it is clear how to extend the definitions given to the case of sequences of more than two phonemes; in order to get the M-rules one has to find the prime implicants of $\bar{\mathbb{B}}^{(k)}$ with $k \geq 3$ (k is the length of the sequence), i.e. the prime implicants of a Boolean function of kn variables. The practical difficulty of this task for larger values of k and n should not be underestimated and here probably further research is necessary. In principle, however, all the M-rules of a language can be computed given only the set of all admissible phoneme sequences of this language (each phoneme being fully specified); furthermore, the M-rules given in the literature, e.g. in HALLE [3] , can be thus tested for compatibility with each other and for conformity with the occurring phoneme sequences.

4. Lexicon rules.

Having computed a set of P-rules (or M-rules) predicting the specifications of certain features the rules can be used to remove these "redundant" specifications from the phonemes. It is common in linguistic practice to replace redundant specifications by blanks. In the set-theoretic notation used here complete removal of redundant elements αr from the phonemes seems to be more adequate. In this section some ideas will be given how to generate the remaining "non-redundant" subsets of phonemes, i.e. the lexicon, by a set of new rules called <u>lexicon rules</u> (L-rules).

According to Sec.3 it suffices to consider the case of P-rules. Let $r \in \mathbb{F}$ be a feature and let

(7) $$a_1 \to \alpha_1 r, \ldots, a_k \to \alpha_k r$$

be k P-rules specifying r in different environments; no a_j contains feature r. Using Boolean notation each $\hat{a}_j$ (j=1,...,k) corresponding to the set a_j is a conjunction of some of the Boolean variables $f_1, \ldots, f_n$ and each of the P-rules is a logical implication saying that the value of the variable r on the right hand side of the rule has a certain unique value if the left hand side has the value true. The value of r is, therefore, predictable if at least one of the conjunctions a_j has the value true, i.e. if the Boolean expression

(8) $$\hat{a}_1 \vee \hat{a}_2 \vee \ldots \vee \hat{a}_k$$

has the value true. The value of r is, therefore, not predictable (i.e. the specification of r can be either + or -) if the Boolean expression

(9) $$\hat{a}_r = \neg \hat{a}_1 \wedge \neg \hat{a}_2 \wedge \ldots \wedge \neg \hat{a}_k$$

which is the negation of (8) has the value true.

This makes possible the formulation of a rule (using again set-theoretic notation)

(10) $$a_r \longrightarrow \pm r$$

called lexicon rule (L-rule) for $\mathbb{B}$, to be read as "if a_r is a subset of some phoneme of $\mathbb{B}$ then both $a_r \cup \{+r\}$ and $a_r \cup \{-r\}$ are subsets of phonemes of $\mathbb{B}$".

These L-rules can then be used to generate the non-redundant phonemic sets of $\mathbb{B}$, starting with the empty set $\emptyset$, by the following prescription: if a phonemic set b occurring in this process of generation contains a_r then it is replaced by the two new phonemic sets $b \cup \{+r\}$ and $b \cup \{-r\}$.

In order to make this process straightforward some additional conventions are introduced:

1. By the usual submatrix interpretation of rule application L-rule (10) is applicable to b if and only if a_r is a subset of b. To test b for applicability of (10) it is, therefore, useful to have b already non-redundantly specified in all the features occurring in a_r in order to avoid having (10) not applicable to b only because the specification of b in one of these features has not yet been filled in. The simplest way of thus ensuring applicability, whenever possible at all, is to introduce an ordering relation $<$ among the features of $\mathbb{F}$ such that for $f,g \in \mathbb{F}$ either $f<g$ or $g<f$ holds. Ordering of the features is quite common in phonology though it is usually introduced at an earlier stage than here. In every prime implicant h of $\overline{\mathbb{B}}$ there is, then, one feature f with the highest rank according to this ordering and we can require that from all the P-rules derivable from h only the single rule having f on the right hand side shall be chosen. Since by Theorem 1.2 h is uniquely determined by each of its P-rules no generality is lost by this special selection. In every P-rule $a_j \longrightarrow \alpha_j r$ of (7) the left hand side then contains only features of a rank less than r, and thus also a_r in (10) contains only features of a rank less than r. All the L-rules are then ordered in a natural way: they are applied in the order of their right hand sides, and the non-redundant specifications are thus filled in "from top to bottom" starting with the lowest-order feature and ending with the highest-order feature.

2. Since any b occurring in the process of generation

contains only non-redundant specifications all specifications in a_r predictable via the P-rules from other specifications of a_r must be removed from a_r. If a_r consists of a single specified feature occurring in a P-rule $c \rightarrow a_r$ then $c \rightarrow \pm r$ is also an L-rule.

3. If $\hat{a}_r$ for some r is always false then the specification of r is always predictable and no L-rule concerning r exists; if, on the other hand, $\hat{a}_r$ is always true or if - which is the case e.g. with the lowest-order feature - no P-rule concerning r exists then the specification of r is never predictable which is expressed by the L-rule $\emptyset \rightarrow \pm r$.

With these additional conventions a set of L-rules is computed by (9) from the P-rules such that for each L-rule there is at least one phonemic set to which it is applicable.

As an example consider the P-rules for tab.1. For the order of features as in tab.1 they are given by the first column in eq.(5). For the lowest-order feature strid there is the L-rule $\emptyset \rightarrow \pm$strid since for this order no P-rule concerning strid exists. For nas there is only the P-rule +strid → -nas, thus -strid → ±nas is an L-rule. The feature cont is predictable from -strid or from +strid (or from +nas), i.e. it is always predictable and no L-rule concerning cont exists. The feature voiced occurs only in the P-rule +nas → +voiced, thus -nas → ±voiced is an L-rule, and since +strid → -nas is a P-rule we get the additional L-rule +strid → ±voiced. Since all a_r consist of only one specified feature no further redundancies have to be removed. Starting from $\emptyset$, application of these rules gives tab.2 containing the lexicon forms of the five phonemes (with respect to the four features). It can be directly verified that tab.2 is filled up by the rules of eq.(5), first column, to give the complete phonemes of tab.1.

For another order of the features one has a different set of P-rules and, consequently, of L-rules. For the n! different orders of the n features there are n! different sets of L-rules each of which gives a different set of lexicon segments (or lexicon matrices). Each of these sets is then filled up by the corresponding set of P-rules to give the complete set of fully specified phonemes (or matrices).

	v	f	m	b	p
strid	+	+	−	−	−
nas			+	−	−
cont					
voiced	+	−		+	−

tab.2

5. Matrices with blanks.

5.1. Distinctness.

As noted in Sec.4 P-rules can be used to remove redundant specifications from phonemes: if $a \rightarrow \alpha r$ is a P-rule and a is a subset of a phoneme $B \in \mathbf{B}$ then the element αr is removed from B. Instead of removing the element αr from B, thus removing the feature r altogether, a common practice in linguistics is to leave the feature r in B but to change its specification into a blank (or zero). It is stressed very much in the literature (e.g. [2] ,p.410) that this blank is not a specification like + or –. This circumstance has been underlined by introduction of the notion of <u>distinctness</u> of phonemes (or phonemic matrices) – see [2] ,p.408. Two phonemes B and C of $\mathbb{B}$ are called distinct if and only if there is at least one feature f $\in \mathbb{F}$ such that B is in f specified as + and C as – (or vice versa); conversely, B and C are said to be not distinct if and only if for every f $\in \mathbb{F}$ either the specification of f in B is identical to that in C or one of both specifications is blank. There has been some discussion about this concept of distinctness (see e.g. [2] ,p.408 f.) and it has been argued that it is not completely sensible; for the present, however, we will accept it as existing and turn to the question "When do phonological matrices remain distinct after the introduction of blanks?".

As before, we consider only the case of single phonemes. Fully specified phonemes are, of course, distinct but they do not necessarily remain so after the introduction of blanks. Taking tab.1 and its P-rules, e.g.(5), as an example the three rules +nas $\longrightarrow$ {-strid,-cont,+voiced} – the right hand sides have been combined for abbreviation – applied to /m/ leave the phonemic set {+nas} or, using the symbol 0, the set {0strid,+nas,0cont,0voiced} whereas the two rules +cont $\longrightarrow$ {+strid,-nas} applied to /f/ leave {+cont, -voiced} or {0strid,0nas,+cont,-voiced} which is not distinct from the result for /m/.

It is possible, however, to have the phonemes of $\mathbb{B}$ pairwise distinct after the introduction of blanks if (as already in Sec.4) an ordering of the features is introduced and if of all the P-rules derivable from a prime implicant of $\overline{\mathbb{B}}$ only the single one with the feature of highest rank on the right hand side is chosen:

Theorem 3

Let the features of $\mathbb{F}$ be totally ordered by an ordering relation $<$ (i.e. for $f, g \in \mathbb{F}$ either $f < g$ or $g < f$) and let $r_j < r$ for all r_j occurring within the left hand side a of a P_rule $a \rightarrow \alpha r$. Then the phonemes of $\mathbb{B}$ are pairwise distinct after introduction of blanks.

Proof:

Let B and C be two (fully specified) phonemes of $\mathbb{B}$, $B \neq C$. Then there is a certain number of features of $\mathbb{F}$ (at least one) such that B is specified contrary to C in exactly these "distinguishing" features and identical to C in the remaining features. Let f be that of the distinguishing features with the lowest rank. Then there is no P-rule $a \rightarrow \alpha f$ which is applicable to both B and C: by assumption all the features in a are of lower rank than f, thus B and C coincide in all features of a. Since the rule is assumed to be applicable to both B and C, a is a subset of both B and C, and since B and C differ in f the set $a \cup \{\bar{\alpha} f\}$ is a subset of either B or C, whatever α. Thus this set is not an implicant of $\overline{\mathbb{B}}$ and, therefore, $a \rightarrow \alpha f$ cannot be a P-rule. This means that no blank can occur on f in B and C, i.e. B and C remain distinct even after the introduction of blanks.

Without ordering of the features two phonemes can become not distinct as is shown by the examples above. Ordering of the features is, however, only sufficient for pairwise distinctness, not necessary, i.e. a set of phonemes with blanks can remain pairwise distinct even without ordering of the features. For an example take the set

$$(11) \quad \begin{cases} +\text{nas} \rightarrow -\text{strid},\ -\text{strid} \rightarrow -\text{cont},\ +\text{strid} \rightarrow +\text{cont}, \\ +\text{cont} \rightarrow -\text{nas},\ +\text{nas} \rightarrow +\text{voiced} \end{cases}$$

of P-rules for tab.1. (One P-rule has been chosen for each of the prime implicants of eq.(4)). This set is not compatible with any ordering of the features since it would require nas $<$ strid, strid $<$ cont and cont $<$ nas which is impossible. Applied to tab.1 in the order given in eq.(11) we get tab.3 with pairwise distinct phonemes.

	p	b	m	f	v
strid	–	–	0	+	+
nas	–	–	+	–	–
cont	0	0	–	0	0
voiced	–	+	0	–	+

tab.3

Unfortunately, there does not seem to be a simple and general necessary condition for pairwise distinctness of phonemes with blanks.

5.2. Uniqueness.

The result of tab.3 depends on the order in (11) of the P-rules. The same P-rules, applied in the order

$$(12) \quad \begin{cases} -\text{strid} \rightarrow -\text{cont},\ +\text{nas} \rightarrow -\text{strid},\ +\text{cont} \rightarrow -\text{nas}, \\ +\text{strid} \rightarrow +\text{cont},\ +\text{nas} \rightarrow +\text{voiced} \end{cases}$$

give tab.4 which is different from tab.3.

	p	b	m	f	v
strid	−	−	0	+	+
nas	−	−	+	0	0
cont	0	0	0	0	0
voiced	−	+	0	−	+

tab.4

In other words, the phonemes with blanks (or, for M-rules, the matrices containing blanks) are not uniquely determined by the P-rules alone but also by the order in which the P-rules are applied to put in blanks.

This situation can be described as follows: Let $P_1 = a \rightarrow \alpha r$ with $a = \{\alpha_1 r_1, \ldots, \alpha_k r_k\}$ be a P-rule which is applicable to a phoneme $B \in \mathbb{B}$, i.e. a is a subset of B. P_1 can then be used to put a blank on r in B. This is, however, impossible if there is already a blank in B on one of the features of a because then a is no longer a subset of B. This blank on a feature r_i of a $(1 \leq i \leq k)$ can be caused only by a P-rule $P_2 = b \rightarrow \alpha_i r_i$ which was applied before P_1. Thus the position of blanks can - and indeed sometimes does, as the examples show - depend on the order of application of the P-rules.

This order dependence somewhat complicates the situation and one can look for ways to avoid it. One way is to give up the submatrix criterion for rule application and to use the non-distinctness criterion instead. Then, the blank on r_i in B would leave a not distinct from B, and P_1 would remain applicable. A serious drawback of this solution is, however, that a blank does not tell which of the specifications + and - has been removed by it. Thus, P_1 would be applicable also to a phoneme B with a blank on r_i which in its full form has

r_i specified as $\bar{\alpha}_i$ instead of α_i. Thus the non-distinctness criterion alone is useless for rule application; it must be amended by criteria ensuring the correct specification of B on r_i.

There is, however, uniqueness even if we use the submatrix criterion for rule application if the features are, as before, totally ordered and only the special P-rules are chosen. For this case we have:

Theorem 4

Let $P_1 = a \rightarrow \alpha r$ with $a = \{\alpha_1 r_1, \ldots, \alpha_k r_k\}$ and $P_2 = b \rightarrow \alpha_i r_i$ be two P-rules applicable to the same phoneme $B \in \mathbb{B}$. Let $r_j < r$ and $s < r_i$ for each s occurring in b. Then there is a further P-rule $P_3 = z \rightarrow \alpha r$ applicable to B such that r_i does not occur in z.

Proof:

> The P-rules P_1 and P_2 are derived from the prime implicants $q_1 = a \cup \{\bar{\alpha} r\}$ and $q_2 = b \cup \{\bar{\alpha}_i r_i\}$ of $\bar{\mathbb{B}}$. Since P_1 and P_2 are both applicable to B both a and b are subsets of B; this means, in particular, that a and b are identically specified in features common to both. Thus, $a \cup b$ is a phonemic set and a subset of B. Let $c = a \setminus \{\alpha_i r_i\}$ and let $h_0 = c \cup b \cup \{\bar{\alpha} r\}$; since $f < r$ for all features f occurring in c or in b, h_0 is a phonemic set. Now h_0 is an implicant of $\bar{\mathbb{B}}$: the feature r_i occurs neither in c (by definition) nor in b (since all features of b are of a rank less than r_i), thus r_i does not occur in h_0. If $h_0 \subseteq C$ for some phoneme C of $\mathbb{B}$ then either $\alpha_i r_i \in C$ or $\bar{\alpha}_i r_i \in C$; therefore, either $c \cup \{\alpha_i r_i\} \cup \{\bar{\alpha} r\} = q_1 \subseteq C$ or $b \cup \{\bar{\alpha}_i r_i\} = q_2 \subseteq C$ which is both impossible since

q_1 and q_2 are implicants of $\overline{\mathbb{B}}$.

If h_o is a prime implicant of $\overline{\mathbb{B}}$ then we have the P-rule $P_3 = z_o \rightarrow \alpha r$ with $z_o = c \cup b$ putting a blank on r, and P_3 is applicable to B since $(c \cup b) \subset (a \cup b) \subseteq B$. If h_o is not prime then there is a proper subset h_1 of h_o such that h_1 is an implicant of $\overline{\mathbb{B}}$. Since each subset of z_o is a subset of B necessarily $\bar{\alpha} r \in h_1$, i.e. $h_1 = z_1 \cup \{\bar{\alpha} r\}$ with $z_1 \subset z_o$. If h_1 is prime then $P_3 = z_1 \rightarrow \alpha r$ is the required P-rule; if not then there is an implicant h_2 of $\overline{\mathbb{B}}$ with $h_2 \subset h_1$ and, similar as before, $h_2 = z_2 \cup \{\bar{\alpha} r\}$ with $z_2 \subset z_1$. Thus we get a sequence of implicants h_i of $\overline{\mathbb{B}}$ with $h_i = z_i \cup \{\bar{\alpha} r\}$ and $z_i \subset z_{i-1} \subset \dots \subset z_o = b \cup c$. If one of the h_i is prime then the sequence terminates with the P-rule $P_3 = h_i \rightarrow \alpha r$. Since the z_i become smaller and smaller the sequence terminates in any case with $h = z \cup \{\bar{\alpha} r\}$ and $z = \{\gamma t\}$, $\gamma \in \mathbb{A}$, $t \in \mathbb{F}$, and h is prime since $z \subset z_o \subseteq B$ and $\{\bar{\alpha} r\} \subset q_1$ and q_1 is prime, thus there is a $C \in \mathbb{B}$ such that $\{\bar{\alpha} r\} \subset C$.

Thus, even if the blank on r in B cannot be put there by P_1 because it is "blocked" by P_2 there is always P_3 which cannot be blocked by P_2 and which puts the blank on r in B. Thus, the position of blanks in the phonemes of $\mathbb{B}$ is uniquely determined by the P-rules alone independent of the order in which they are applied.

Literature.

[1] R.JAKOBSON and M.HALLE: Fundamentals of Language. Den Haag, 1956.

[2] R.STANLEY: Redundancy Rules in Phonology. Language 43, 393 (1967).

[3] M.HALLE: The Sound Pattern of Russian. 's Gravenhage, 1959.

[4] G.UNGEHEUER: Studia Linguistica XIII, 60 (1959).

[5] N.CHOMSKY and M.HALLE: The Sound Pattern of English. New York, 1968.

[6] I.B.PYNE and E.J.McCLUSKEY,Jr.: The Reduction of Redundancy in Solving Prime Implicant Tables. IRE Trans. EC-11, 4 (473), Aug. 1962.

[7] M.HALLE: On the Bases of Phonology. In: J.A. FODOR and J.J.KATZ (eds.): The Structure of Language, Englewood Cliffs, N.J., 1964.