SYNTHESIS BY RULE AS A TOOL FOR PHONOLOGICAL RESEARCH

IGNATIUS G. MATTINGLY
Haskins Laboratories and University of Connecticut

SYNTHESIS BY RULE AS A TOOL FOR PHONOLOGICAL RESEARCH*

IGNATIUS G. MATTINGLY

Haskins Laboratories and University of Connecticut

The phonologist, having analysed a corpus of phonetic data and proposed tentative phonological rules, must test these rules by generating novel utterances for acceptance or rejection by an informant. There are certain methodological problems in such testing which could be avoided by the use of speech synthesized' by rule. Synthesis by rule is now a practical technique, and a synthesis-by-rule system at Haskins Laboratories has been used to formulate rules for General American English.

A grammar, as Chomsky (1965: 3-9) has reminded us, is a theory of a particular language, consistent, on the one hand, with the observable structures of utterances in the language, and on the other hand, with some higher-level theory explaining language as such and incorporating known linguistic universals. The concern of the phonologist is to construct the portion of the grammar of the language which gives the rules of its sound system. Like other linguists, he goes about his task by eliciting a corpus of utterances from a native informant and formulating tentative rules which account for this corpus. He then produces new utterances according to these rules, which he submits to his informant. If the informant considers these utterances acceptable, the rules are, for the time being, corroborated; if the informant corrects any of the utterances, the rules are, at least in part, falsified[1] and the phonologist must revise his rules in the light of the informant's corrections. This process of rule formulation and revision continues in principle until a set of phonological rules has been constructed which (in conjunction with other parts of the grammar) predict all possible utterances of the language, and no others; it continues in practice until the phonologist feels he can no longer readily improve upon his current version of the rules.

According to this account of the phonologist's task, the crucial stage is the generation of novel utterances according to the proposed rules and their criticism by the informant. Whether the rules are arrived at by some rigorous analytic procedure or by a series of inspired insights is irrelevant; whether the rules correctly generate novel utterances is vital.

But it is at just this stage that various difficulties embarrass the phonologist. How can he be certain that the novel utterances he generates are what he takes them to be: true products of his rules? Though he does in fact know the correct rule, he may accidentally generate an utterance which violates the rule. If his informant corrects such an utterance, the phonologist will probably then realize that the error was in his performance rather than in his rules. A more dangerous possibility is that he will, being a learner as well as an analyst of the language, unconsciously acquire and

[1] *The notions of falsification and corroboration derive from Popper (1959).*

correctly apply a rule which has not come to his conscious attention as a phonologist. He is quite likely to overlook such a rule completely. A similar fate may befall rules which are common to the phonologist's native language and the language he is studying. Again, there may be points about which the informant may vacillate. Here a formal test is required, ideally with a carefully controlled set of stimuli. Yet the linguist may find the production by natural means of a satisfactory set of stimuli a difficult matter. Finally, even supposing that the test utterances follow the rules, and only the rules, there is the possibility of bias in the selection of the utterances. If the informant could, somehow, use the rules to generate utterances of his own selection, the rules would be put to a much more severe test. But this kind of test seems to require that the informant play the role of linguist, and hitherto phonological descriptions have been tested in this way only in the case of those relatively few languages of which linguists are native speakers—and usually the test has taken the form of a critical review of already published work.

I would like to suggest that the techniques of speech synthesis by rule could do much to relieve the phonologist of such embarrassments. Imagine an automatic system, the inputs to which are proposed phonological rules of a language, and a phonemic transcription of an utterance of the language, and the output from which is a synthetic acoustic representation of the utterance. Such a system simulates the phonologist in his generative phase. But it does not make accidental errors, and it applies only rules which have been explicitly stated. A native informant can propose utterances to be generated—or even, if he has learned the transcription system, generate the utterances himself—and report to the phonologist in what respects the synthetic versions are incorrect. In difficult cases, the informant can be invited to compare stimuli produced by alternative versions of the rules differing only with respect to the variable of interest. In the light of the informant's responses, the rules can be revised easily and quickly, and the informant can then be confronted with the output of the revised rules.

Much of what we have just been saying has already been said by Lisker *et al.* (1962). These authors stressed the importance of the generative phase of the phonologist's work, the valuable role of synthetic speech in phonological investigation, and the advantages of using synthesis by rule for making phonological statements. They and their colleagues at Haskins Laboratories had even compiled a set of rules for English (Liberman *et al.*, 1959). But at the time they wrote the technology essential to practical synthesis by rule was not yet far advanced. Indeed, though synthesis by rule has been a research objective since the time of von Kempelen (1791), only quite recently has it reached the stage where it can be really useful to the phonologist.

A number of problems had to be solved first, some of them technical, others phonetic and phonological. Reliable synthesizers, capable of simulating speech reasonably well, had to be developed. The cumbersome mechanical synthesizers of the 18th and 19th centuries were succeeded in the 20th century by electrical synthesizers, like the Voder (Dudley *et al.*, 1939), the Haskins Pattern Playback (Cooper, 1950), and PAT (Lawrence, 1953), which had more predictable characteristics and were easier to modify. Advances in electronics during the past ten years have at last made possible the really

reliable performance of the current generation of synthesizers: OVE III (Liljencrants, 1968), and the Haskins (Mattingly, 1968a, b) and Glace-Holmes synthesizers (Glace, 1968).

There was also the question of the extent to which it is practical to simulate the process of speech production in the design of a synthesizer. Von Kempelen attempted to build a complete vocal tract analogue—a goal which has yet to be realized satisfactorily. The Voder and the Pattern Playback, on the other hand, merely represented the acoustic spectrum of speech. PAT incorporated more specific assumtions about the speech signal: the formants of speech are represented in its design by resonant circuits; the frequencies of the resonances and the character of their excitation are parameters of synthesis whose momentary values are determined by control functions. Resonance synthesizers like PAT proved to be very practical, and have been used for most of the research in synthesis by rule until the recent renewal of interest in vocal tract analogues (Coker, 1968; Matsui *et al.*, 1969; Nakata *et al.*, 1968).

A convenient method of dynamically controlling a synthesizer had also to be devised. Synthesizers as late as the Voder were manually operated: the 'rules' were part of the operator's skill. The Playback, however, was controlled by painted patterns on a moving plastic belt—a great advance because the rules could now be explicitly stated in terms of these control patterns (Liberman *et al.*, 1959). In the case of PAT, parametric functions are drawn on a similar belt, and the rules stated in terms of the functions. Drawing patterns or functions to rule is a time-consuming activity, however, and little speech was actually synthesized by rule until it became economic to use a digital computer not only to transmit the control functions to the synthesizer but also to calculate these functions (Kelly and Gerstman, 1961).

Besides overcoming these technical problems, it was necessary to achieve an understanding of the relationship between speech at the acoustic level and speech at the phonetic or perceptual level, that is, the output of the phonology. This relationship, as is now widely realized, is anything but straightforward: speech is not simply a train of distinct and separately perceived acoustic elements, each corresponding to a symbol of the phonetician's narrow transcription. In reconstructing the speaker's message, the listener relies on an elaborate system of cues which are not necessarily characterized in an acoustically obvious way and which overlap in time. The relationship, now fairly well understood after many years of research in acoustic phonetics and speech perception, has been aptly called the "speech code" by Liberman *et al.* (1967). Obviously, a system for testing phonological rules must incoporate a reliable equivalent of the speech code; in other words, a sub-system of general phonetics which provides an accurate acoustic interpretation of the phonologist's narrow transcription. Otherwise, there will be doubt whether an error is phonological or phonetic, and at worst, the synthetic speech will be unintelligible and no evaluation of the phonology can be made. While further progress is necessary, not only in understanding the speech code but also in simulating it, enough is known about the code so that intelligible speech can be synthesized by rule and phonological errors can be readily detected.

Such speech is not very natural-sounding. The reason is that our knowledge of the speech code, like our knowledge of anything in nature, is "minimal" (Liberman *et al.*, 1959): we know only the most important facts. It is probably the case that only the most powerful cues for the various phonetic distinctions can be experimentally isolated. The "naturalness" of the natural speaker, however, consists in part in his casualness about these powerful cues; he is apt to omit them or represent them defectively in his performance. He gets by with this because he can count on the combined effect of a host of minor cues, each in itself of little weight. Synthetic speech cannot afford to be so casual. If we bias our phonetic rules in favour of naturalness (insofar as we know how to do this), we lose intelligibility. Fortunately, we are on safe ground for purposes of phonological research as long as the most powerful cues are properly represented. Experience thus far indicates that as long as the synthesized sounds are accepted as speech at all, lack of naturalness does not interfere with the perception of phonological rules. And it would even seem preferable to base a phonology on a general phonetics which is minimal, if it can thereby be well specified. The linguist who uses his own natural speech has a richer system, but one that is variable and poorly defined.

A further difficulty arises because no synthesizer yet built straightforwardly simulates all the acoustic characteristics of speech; various engineering compromises must inevitably be made. Hence it is not enough to know what the cues are; what Fant *et al.* (1962) have called "synthesis strategy" is necessary: a set of procedures for producing the most acceptable acoustic images of the various cue-carrying elements of speech within the limitations of the design of a particular synthesizer. While synthesis strategy will always be a problem, it has shrunk to quite manageable dimensions for the current generation of synthesizers. Thus, while it is difficult to produce on any synthesizer a stop consonant burst which bears a close acoustic resemblance to its natural prototype, it is not so difficult to produce a sound which is perceptually acceptable as a burst, and hence the formulation and testing of phonological rules relating to exploded and unexploded stops can proceed.

Finally, there must be a suitable format for the statement of phonological rules. The format must be restrictive enough for its interpretation to be manageable by a computer programme which applies the rules, yet flexible enough to enable any rule to be stated, and to allow rules to be relatively ordered. While a completely general format has yet to be devised, very encouraging progress has been made, and it is becoming fairly obvious what further steps need to be taken.

None of these problems, of course, even exists for the phonologist using conventional methods. Since he is a human being, he has a perfect synthesizer, his own vocal tract, and a neural system for the encoding and decoding of speech which he can take more or less for granted. He can state his phonological rules in any form he likes. But, as we have seen, his methods have their own shortcomings, and since the most crucial problems associated with synthesis by rule have been largely overcome, it is now time for him to consider whether the technique could not alleviate some of these shortcomings.

As an illustration of the use of synthesis by rule for phonological research, we will describe the system at Haskins Laboratories, which has served for some time as a tool for the investigation of Southern British English and General American English (Haggard and Mattingly, 1968; Mattingly, 1968a, b). The system consists of a parallel resonance synthesizer, an appropriately programmed Honeywell DDP-224 computer and conventional peripheral devices. The synthesizer (designed by Franklin S. Cooper and Robert Epstein), has four resonant circuits, representing the first four formants of the speech signal. Vocal-cord excitation is simulated by a buzz generator; its frequency is the fundamental frequency of the synthetic speech. Noisy excitation is simulated by a hiss generator. By setting buzz and hiss switches it is possible to excite the resonances with buzz, with hiss, with both or with neither. There are also three high-pass filters which shape noise for the sounds [s], [ʃ] and [f, θ]. Depending on the setting of the fricative switches, the hiss generator excites one of these three filters, or none of them. The amplitude of the excitation of each of the resonances and filters is regulated by a modulator. Their outputs are added together in a mixer and the amplitude of the output signal is regulated by a variable-gain amplifier.

The dynamically controllable parameters of the synthesizer are the states of the buzz and hiss switches, hereafter considered as a single parameter BH with four possible values; the frequency F0 of the buzz generator; the amplitudes M1, M2, M3 and frequencies F1, F2, F3 of the first three formant resonances; the state of the fricative selection switch FS; the amplitude MF of the fricative excitation; and the setting OA of the variable-gain amplifier. Coded values for these parameters are loaded by the computer into two 24-bit registers, from which the various circuits of the synthesizer are controlled. For an utterance to be synthesized, it must be represented in the computer memory by a series of pairs of 24-bit words, each pair containing a set of parameter values. During synthesis, the computer transmits one parameter value set after another, at 5 msec. intervals, to the two control registers; the state of the synthesizer is determined by the most recently transmitted set.

The calculation of parameter value sets for an utterance is the chief task of the synthesis by rule programme.[2] The inputs to the programme are a set of rule tables for some dialect, and a string of symbols, such as the following:

H EE — "T OO K — A — 'W AW K — "E V R EE — 'M AW R N I NG.

Upper case letters and digraphs represent phonemes (this notation is not only convenient for input through the computer typewriter but allows us to distinguish

---

[2] *The programme is based on two earlier programmes developed at J.S.R.U., Eastcote: one for the synthesis of segmental phonemes (Holmes et al., 1964); the other for the synthesis of prosodic features (Mattingly, 1966). The functions of these two programmes have been combined, and various refinements have been made: non-linear formant frequency transitions, the machinery for the statement and application of allophone rules, a more elaborate procedure for the specification of excitation, generalization of the method of calculating F0, and immediate playback of the synthetic speech.*

synthetic phonemes from "real" ones). A juncture sign serves as a context marker for the phonological rules which deal with word-boundary phenomena. Three intonation contours—fall, rise and fall-rise—can be indicated, and the intonation signs also mark the pause which follows the end of a breath-group. Three degrees of lexical stress—major, minor and low—can be indicated, as can prominence, i.e. the change— usually a rise—in F0 which occurs on major-stressed syllables of one or more words in a sentence under certain syntactic and semantic conditions. The experimenter types this input string on the computer typewriter, the parameter value sets are immediately calculated, and the synthetic utterance can then be heard.

The rule tables, which contain in coded form the information necessary for the synthesis of phonemes and prosodic features, fall into three main parts: the pitch table, a set of variables used to calculate F0; a set of segmental phoneme tables containing the information needed to calculate the other parameters; and a set of allophone rules, each of which contains a statement of the phoneme classification and the context to which the rule applies, and the change in the segmental table required to implement the rule. The phoneme classification is stated in terms of a few elementary binary distinctive features; any sub-set of the inventory definable by these features can be specified. Similarly, the context is described as a combination of elementary binary contextual features such as "prevocalic," "postvocalic," "prejunctural," and so on. The table change is a modification of one of its entries by incrementing, decrementing or multiplying by a specified quantity, or substituting a specified quantity, or replacement by another table entry.

During the first pass of the programme, as the input string is read in, each phoneme is classified according to its distinctive and contextual features. During the second pass, before the calculation of the parameter values for a phoneme or its neighbours takes place, this classification is compared with each of the allophone rules to determine which rules are relevant, and the appropriate modifications are made in the segmental table for that particular occurrence of the phoneme. Since the rules are tested for relevance and applied in the order in which they are stated, the ultimate effect on the phoneme tables may depend on the particular order of rules.

The sequence of modified segmental tables contains the information required to calculate the values for parameters other than F0. On the basis of this information the duration for each phoneme and the steady-state values for fricative and formant amplitudes and formant frequencies are assigned. An initial and final transition for each parameter is computed; their beginning and end points and their duration can be controlled by various table entries so that transitions can be constructed appropriate to the manner class and point of articulation of the phoneme and to those of its neighbours. An excitation sequence is determined for the phoneme, consisting of periods of buzz, hiss, mixed buzz and hiss or silence, in appropriate order. A period of closure voicing (as in voiced stops) and a period of hiss at the beginning of the following phoneme (to represent aspiration) can be specified. Similarly, a frication sequence is determined.

During the third and final pass of the programme, F0 is calculated with the help

of the pitch table. The calculation is carried out syllable by syllable, the boundaries of a syllable being defined by an arbitrary rule. For the part of the sentence up to the tonic syllables, values assigned to variables in the pitch table specify changes in the current value of the pitch and the slope of the pitch function for the syllable, depending on whether the syllable is prominent or, if not prominent, whether it follows a prominent syllable; and on whether a voiceless interval precedes the voiced part of the syllable. Other variables in the pitch table specify the shape of the intonation contour specified in the input sentence and imposed on the tonic syllable and any following syllables.

The system just described is not a *model* of the phonological process. For the system to be a model, it would have to be part of a larger system which modelled the entire grammar of the language, and included a lexical and syntactic component; the rules would have to be stated in a way which parallels their statement in the speaker's neuromuscular system; the synthesizer would have to be a true vocal tract analogue, capable of producing all and only those sounds which the human vocal tract produces, and consequently having a very different set of parameters—to mention only a few of the most important requirements.[3] The much more modest claim is made that the system is a *tool* for phonological research. The correspondence of the system and the rules to theoretical phonology and phonetics is quite sufficient for the purpose of stating a body of phonological rules and producing synthetic utterances which conform to these rules.

The synthesis by rule system—i.e. the programmed computer and the synthesizer—correspond, more or less, to a general phonetics. Given the information in a succession of segmental tables, the system produces the corresponding utterance. Such a system is, in principle, capable of dealing with any language (or, more precisely, any dialect) described by a set of rules in the prescribed format. It takes care of the physical universals of language, linking successive steady states of formant frequencies and amplitudes by appropriate transitions, and providing for what Wang and Filmore (1961) have called "intrinsic" allophones: the variations in acoustic realization dependent on phonetic context which are the same in all languages because they result simply from articulatory mechanics. The system produces the required excitation, and calculates a pitch contour for each breath-group, subject to the rules for stress and intonation. In practice, of course, the particular system we have described is not complete: for example, the programme has no provision for dealing with tone languages and the limited fricative repertoire of the synthesizer would make production of certain sounds in some languages difficult. But deficiencies of this kind can be readily remedied and do not invalidate the general usefulness of the system.

The succession of segmental tables, modified by the allophone rules, corresponds to the phonologist's narrow transcription of an utterance. These modified tables contain the language-dependent and utterance-dependent information needed to control the general phonetics. Of course, in their present form, the tables contain instructions for calculating acoustic parameters; they might more consistently be replaced by a

---

[3] *For a stimulating discussion of synthesis by rule as a model, see Kim (1966).*

specification of general phonetic features, with the translation to parameter instructions being carried out by the programme itself. The explanation for this inelegance, is that, as a practical matter, we wished to be free to make adjustments in two areas already referred to: the speech code and the synthesis strategy. Both of these functions were most conveniently accomplished by incorporating the information controlling the parametric translation in the tables, for the time being. But the presence of this information is hardly an impediment to phonological investigation.

The inventory of phonemes for which segmental tables are given in the rules, the allophone rules for modification of segmental tables in context, the phonetic characterization given for each phoneme in the tables, and (in the case of prosodic features) the pitch tables, correspond to the phonology of the dialect. The statement and ordering of the allophone rules are clearly the central task of the phonologist; the unmodified segmental table simply describes the residual or unmarked allophone, to be used if none of the allophone rules apply to a phoneme. Similarly, the pitch table may be regarded as consisting of certain entries which characterize the pitch line of the unmarked breath-group in the absence of prominence and of certain other entries which modify the pitch line as required by the occurrence of phonological structures such as prominence or marked intonation contours.

The statement of an allophone rule requires a determination of what is best regarded as the unmarked allophone, of the context which conditions a change from the unmarked case, of the nature of this change, and of the alterations in the table required to effect this change. This sounds straightforward enough until one considers, for instance, how to state the General American rule for strongly aspirated voiceless stops. The rule is given by Bloch and Trager (1942: 42-43) for /t/ but applies to /p/ and /k/ as well:

> Aspirated [t'] initially before a vowel (*tin, tomorrow*); medially between vowels, if the following vowel has the louder stress (*attack*); medially after any consonant except [s], before a stressed vowel (*captivity, dictation, entire, particular*, etc.); finally after a vowel or any consonant, but here in free variation with [t⁼] and [t'] (*at, apt, act, cast, raft, wished, melt, ant*, etc.).

Let us ignore the free variation in final position, and postpone, for the moment, consideration of what happens after [s]. In order to state the other facts properly as allophone rules, we have to recognize that the weakly aspirated allophone is unmarked and that in two distinct contexts—post-junctural and pre-stress—two distinct phonological structures happen to be signalled by the same phonetic phenomenon: strong aspiration. An allophone rule for each of these contexts can then readily be formulated. The wrong decision about the unmarked allophone, or a failure to recognize that two distinct phonological structures are involved, would have resulted in incorrect rules or prevented formulation of the rules altogether.

Matters become more complicated when the relative ordering of the rules has to be considered. For instance, voiceless stop bursts are modified by a number of contexts: (1) a following stop (no burst); (2) a following major-stressed syllable (burst

louder than that of the unmarked allophone); (3) a following low-stressed syllable (softer burst); (4) a preceding [s] (no burst). Since contexts (1), (2) and (3) are mutually exclusive, we do not have to worry about the relative order of rules for these contexts. Since the effect of the rules (1) and (4) is the same, we do not have to worry about order in this case, either. But (4) *must* follow (2) and (3), since the [s] rule overrides the stress rules. (Similarly, a rule de-aspirating stops after [s] follows and overrides the stress rules controlling aspiration discussed above.) Also, it is clear that if it were not possible to order the rules at all, we would have to write rules for five rather than three contexts.

The durational rules for vowels illustrate a different aspect of ordering. Four phonological structures come into play here; intrinsic vowel length, stress, the phonotactic rules relating to syllable-final voiced clusters, and the rules for pre-pausal phonemes. It was not found possible to state the effect of stress on vowel length in any regular way, so major-stress and low-stress durations are stored as spare entries in the table for each separate vowel. The first two allophone rules affecting vowel length replace the unmarked (minor-stress) duration by the major—or low—stress durations when appropriate. Then, if a vowel is followed by a voiced cluster and is either major-stressed or minor-stressed, its duration is increased by 1.25. And in a pre-pausal syllable, the duration of a vowel, like that of any other phoneme, is increased by 1.5. Rules of this sort require not only consideration of their ordering—the effect would obviously be quite different if the stress rules came after the voiceless cluster or pre-pausal rules—but also of the cumulative effects of several changes on the same table entry. The scheme described works well for the most part, but in a few instances, where stress, a voiced cluster and a pause have all increased the duration of the vowel, the result is obviously *too* long.

Problems such as we have just been discussing are not, I think, artifacts of the system; they are real problems which must be overcome if a rigorous and complete statement of the phonology of a language is to be achieved. For the reasons we noted earlier, it is quite possible for a linguist to produce natural utterances which satisfy his native informant without having correctly and explicitly worked out the rules; it is also quite possible for a narrative account of phonology to omit or misstate rules without the omission or misstatement being very obvious. But synthesis by rule is a sterner taskmaster. The synthesis of any appreciable amount of text exposes to the phonologist and even more clearly to his informant the inadequacies of the current version of the phonology of the language under study.

As long as the main goal of phonology was considered to be the extraction of the inventory of phonemes, there was some excuse for the shirking of rigorous testing of phonologies. More recently, however, we have come to see that the phonological rules which link the surface structure to the phonetic level are the phonologist's true objective. Thus there now seems no escaping the obligation to subject the rules to the severest possible tests, and the tool essential for such tests, synthesis by rule, is available.

REFERENCES

BLOCH, B. and TRAGER, G. L. (1942). *Outline of Linguistic Analysis* (Baltimore).

CHOMSKY, N. (1965). *Aspects of the Theory of Syntax* (Cambridge, Mass.).

COKER, C. H. (1968). Speech synthesis with a parametric articulatory model. *Preprints Speech Symposium, Kyoto 1968*, A-4.

COOPER, F. S. (1950). Spectrum analysis. *J. acoust. Soc. Amer.*, 22, 761.

DUDLEY, H., RIESZ, R. R. and WATKINS, S. S. A. (1939). A synthetic speaker. *J. Franklin Inst.*, 227, 739.

FANT, C. G. M., MARTONY, J., RENGMAN, U. and RISBERG, A. (1962). OVE II synthesis strategy. *Sp. Commun. Sem.* (Stockholm), F5.

GLACE, D. A. (1968). Parallel resonance synthesizer for speech research. *J. acoust. Soc. Amer.*, 44, 391.

HAGGARD, M. P. and MATTINGLY, I. G. (1968). A simple programme for synthesizing British English. *IEEE Trans. Audio*, 16, 95.

HOLMES, J. N., MATTINGLY, I. G. and SHEARME, J. N. (1964). Speech synthesis by rule. *Language and Speech*, 7, 127.

KELLY, J. L. and GERSTMAN, L. J. (1961). An artificial talker driven from a phonetic input. *J. acoust. Soc. Amer.*, 33, 835.

VON KEMPELEN, W. R. (1791). *Mechanismus der menschlichen Sprache nebst der Beschriebung seiner sprechenden Maschine* (Vienna).

KIM, C-W. (1966). The linguistic specification of speech. *Working Papers in Phonetics*, No. 5, U.C.L.A. (Los Angeles).

LAWRENCE, W. (1953). The synthesis of speech from signals which have a low information rate. In *Communication Theory*, ed. Jackson, W. (London), 460.

LIBERMAN, A. M., LISKER, L., DELATTRE, P. C. and COOPER, F. S. (1959). Minimal rules for synthesizing speech. *J. acoust. Soc. Amer.*, 31, 1490.

LIBERMAN, A. M., COOPER, F. S., SHANKWEILER, D. P. and STUDDERT-KENNEDY, M. (1967). Perception of the speech code. *Psychol. Rev.*, 74, 431.

LILJENCRANTS, J. C. W. A. (1968). The OVE III speech synthesizer. *IEEE Trans. Audio*, 16, 137.

LISKER, L., COOPER, F. S. and LIBERMAN, A. M. (1962). The uses of experiment in language description. *Word*, 18, 82.

MATSUI, E., SUZUKI, T., UMEDA, N. and OMURA, H. (1968). Synthesis of fairy tales using an analog vocal tract. *Reports of the Sixth International Congress on Acoustics*, ed. Kohashi, Y. (Tokyo), B 159.

MATTINGLY, I. G. (1966). Synthesis by rule of prosodic features. *Language and speech*, 9, 1.

MATTINGLY, I. G. (1968a). *Synthesis by rule of General American English*. Supplement to Status Rep. on Sp. Res., Haskins Labs. (New York).

MATTINGLY, I. G. (1968b). Experimental methods for speech synthesis by rule. *IEEE Trans. Audio*, 16, 198.

NAKATA, K., ICHIKAWA, A. and MIURA, T. (1968). Vocal tract analog speech synthesis by rule. *Preprints Speech Symposium, Kyoto 1968*, A-5.

POPPER, K. (1959). *The Logic of Scientific Discovery* (New York).

WANG, W. S-Y. and FILLMORE, C. J. (1961). Intrinsic cues and consonant perception. *J. Sp. Hrg. Res.*, 4, 130.