[Mechanical Translation and Computational Linguistics, vol.9, nos.3 and 4, September and December 1966]

Historical Change in Language Using Monte Carlo Techniques*

by Sheldon Klein, Carnegie Institute of Technology, Pittsburgh, Pennsylvania, and System Development Corporation, Santa Monica, California†

A *system has been programmed in JOVIAL to serve as a vehicle for testing hypotheses about language change through time. A basic requirement of the system is that models must be formulated within the framework of Sapir's concept of drift and Bloomfield's definition of a speech community. Outside these restrictions, an experimenters selection of hypotheses is free. The system, which can be viewed as performing Monte Carlo simulations of group, language change, has been successfully tested in several computer runs using an extremely simple model of linguistic interaction. (The system, and any model tested within its framework, are separate entities. Accordingly, the use of a trivial model to check out the operation of the system does not depreciate its ability to handle models of vast complexity.) The initial test population consisted of fifteen adults and five children, each represented by a phrase-structure generation-recognition grammar. The grammars and the frequency parameters associated with their individual rules were not necessarily identical. During the course of a run some individuals died and others were born. Newborn children acquired the language of the community. The units of interaction consisted of conversations that were produced by the grammars of speakers and parsed by the grammars of auditors. The linguistic structure of a conversation determined changes in the auditor's grammar. Decisions in the system were made with random numbers on the basis of weighted frequency parameters. To insure control of free variables before undertaking experiments with factors causing change, the goal of the initial experiment was to obtain a condition of linguistic stability and essentially identical results for the population as a whole from several computer runs which differed only in the choice of random numbers referred to in decision-making processes. Such results were obtained; even though the fate of individual members of the speech community varied widely in the different trials, the mean values of the frequency of the grammatical rules in the total population were very similar at identical time periods in each run, for a simulated span of twenty-five years and the structure equilibrium state.*

I. Introduction

Computer simulation of real-world events for the purpose of prediction or of testing the validity of models has numerous precedents in the behavioral sciences.[1-8] The first step in such a simulation is the formulation of a model in terms that can be implemented in a computer program. A strong check on the validity of the assumptions in the model is successful prediction of pertinent events. For some types of simulation, such as the behavior of laboratory animals in a hypothetical experiment, a model can be considered adequate if the simulated behavior falls only within the range of behavior of real animals in a live experiment. In general, a model can be considered valid even if its predictions are only statistically significant approximations of real-world behavior.

Simulation experiments may model the behavior of a single entity or that of a large population. The number of entities used in a simulation may be equal to a total population or may be viewed as representing a small sample of a very large population.

The term "Monte Carlo," adopted because of its gambling connotations, refers to the use of random numbers as determiners of events in a simulation. The events that take place may be random only within the constraints of posited stochastic relationships that govern probabilities of transition from one state of events

* This research is supported in part by grant MH-07722, National Institute of Mental Health, U.S. Public Health Service (to the Carnegie Institute of Technology). Portions of this paper were presented at the 1964 and 1965 winter meetings of the Linguistic Society of America and before the Computation and Control Colloquium, Harvard University, March, 1966. The author is grateful to Herbert A. Simon, John T. Gullahorn, and Frank N. Marzocco for their comments and suggestions.

† Now at the University of Wisconsin, Madison.

to another. The transition probabilities may be either constant or altered during the course of a simulation. Assume, for example, that under certain conditions a given event has a 0.2 chance of occurring. Further assume that the pertinent conditions exist. The simulation system would refer to a source of random or pseudorandom numbers for a fraction in the range 0-1, implementing the event only if that number were in the range 0-0.2.

In evaluating the predictions of a system incorporating such decision-making devices, it is essential to determine the effects of different choices of random numbers. This is normally accomplished by repetition of the same simulation with different random numbers. The pertinent data may then appear in the form of a statistical analysis of the behavior in the repeated trials.

A simulation may yield several kinds of information of interest to a researcher. For example, it might be of interest to know that a model predicted a state *C* from a state A and also to know that in the course of prediction it simulated an intermediate state B.

The program described in this paper is a vehicle for the testing of diverse models of language change. While, in the course of my work, I may test the implications of some particular models, the program itself will serve, hopefully, as a general tool for conducting a variety of simulation studies.

II. The Basic Design of the Simulation System

The program, which is written in JOVIAL, an ALGOL compiler language, is designed to simulate the interaction of members of a speech community among themselves and with members of other communities. It is flexible enough to model special relations among particular members, for example, family groups and social classes; to simulate the transmission of language from one generation to the next; and to handle the phenomena of multilanguage acquisition.

While the experimenter has a large range of choice in designing models for simulation, certain basic assumptions about group language phenomena are inherent in the design of the program and are more or less unalterable. Such assumptions are analogous to definitions and metatheorems in a system of formal logic. Except for the concept of "generation grammar," none of these primitive assumptions is alien to readers of Sapir and Bloomfield. The assumptions are consistent with Sapir's concept of "drift" (ref. 9, pp. 165-66):

> Language exists only in so far as it is actually used—spoken and heard, written and read. What significant changes take place in it must exist, to begin with, as individual variations. This is perfectly true, and yet it by no means follows that the general drift of language can be understood* from an exhaustive descriptive study of these variations alone. They themselves are random phenomena,† like the waves of the sea, moving backward and forward in purposeless flux. The linguistic drift has direction. In other words, only those individual variations embody it or carry it which move in a certain direction, just as only certain wave movements in the bay outline the tide. The drift of a language is constituted by the unconscious selection on the part of its speakers of those individual variations that are cumulative in some special direction. This direction may be inferred, in the main, from the past history of the language. In the long run any new feature of the drift becomes part and parcel of the common, accepted speech, but for a long time it may exist as a mere tendency in the speech of a few, perhaps of a despised few. As we look about us and observe current usage, it is not likely to occur to us that our language has a "slope," that the changes of the next few centuries are in a sense prefigured in certain obscure tendencies of the present and that these changes, when consummated, will be seen to be but continuations of changes that have already been effected.

The basic assumptions of the simulation system are also consistent with Bloomfield's thoughts about the nature and formal representation of the concept of "speech-community" (ref. 10, pp. 46-47).

> The most important differences of speech within a community are due to differences in *density of communication.* The infant learns to speak like the people round him, but we must not picture this learning as coming to any particular end: there is no hour or day when we can say that person has finished learning to speak, but, rather, to the end of his life, the speaker keeps on doing the very things which make up infantile language-learning . . . Every speaker's language, except for personal factors which we must here ignore, is a composite result of what he has heard other people say.
>
> Imagine a huge chart with a dot for every speaker in the community, and imagine that every time any speaker uttered a sentence, an arrow were drawn into the chart pointing from his dot to the dot representing each one of his hearers. At the end of a given period of time, say seventy years, this chart would show us the density of communication within the community. Some speakers would turn out to have been in close communication: there would be many arrows from one to the other, and there would be many series of arrows connecting them by way of one, two, or three intermediate speakers. At the other extreme there would be widely separated speakers who had never heard each other speak and were connected only by long chains of arrows through many intermediate speakers. If we wanted to explain the likeness and unlikeness between various speakers in the community, or, what comes to the same thing, to predict the degree of likeness for any two given speakers, our first step would be to count and evaluate the arrows and series of arrows connecting their dots. We shall see in a moment that this would be only the first step; the reader of this book, for instance, is more likely to repeat a speech-form which he has heard, say, from a lecturer of great fame, than one which he has heard from a street-sweeper.

* "Or rather apprehended, for *we* do not, in sober fact, entirely understand it as yet"[ref. 9, p. 166, n. 8].

† "Not ultimately random, of course, only relatively so" [ref. 9, p. 166, n. 9],

The chart we have imagined is impossible of construction. An insurmountable difficulty, and the most important one, would be the factor of time: starting with persons now alive, we should be compelled to put in a dot for every speaker whose voice had ever reached anyone now living, and then a dot for every speaker whom these speakers had ever heard, and so on, back beyond the days of King Alfred the Great, and beyond earliest history, back indefinitely into the primeval dawn of mankind: our speech depends entirely upon the speech of the past.

Since we cannot construct our chart, we depend instead upon the study of indirect results and are forced to resort to hypotheses. We believe that the differences in density of communication within a speech-community are not only personal and individual, but that the community is divided into various systems of sub-groups such that the persons within a sub-group speak much more to each other than to persons outside their sub-group. Viewing the system of arrows as a network, we may say that these sub-groups are separated by *lines of weakness* in this net of oral communication. The lines of weakness and, accordingly, the differences of speech within a speech community are *local*—due to mere geographic separation—and *non-local,* or as we usually say, *social.*

Simulation of drift through a dynamic implementation of Bloomfield's concept of speech community, in which the density of communication is determined by probability values rather than statically mapped by lines of interaction, is a goal implicit in the design of the simulation system. Any programing of models or testing of hypotheses with this program must take place within this basic framework.

A. POPULATION

Each member of a speech community is represented in the program by a generation grammar and a recognition grammar. Individuals with command of more than one language may be associated with additional grammars. A grammar consists of a set of rules for either parsing or generating forms in a particular language.

The grammars of individuals are not necessarily identical. During the course of a simulation, various individuals will die, and new ones will be born. A death requires the deletion of the grammars associated with the deceased; a birth, the addition of new grammars. The grammars representing newborn children are empty. An adult just entering an alien speech community may acquire empty recognition and generation grammars in addition to the non-empty ones he may possess as a member of another speech community.

The program is flexible with respect to the kinds of recognition- and generation-grammar rules it may use. These rules may be limited just to syntax, just to phonology, or to syntax and semantics; or they may pertain to any range of linguistic phenomena that some theory might designate as significant. Accordingly, the program can use either stratificational or transformational grammar models and might manipulate rules pertaining to phonemes or distinctive features, semolexemic rules or transformations.

This flexibility is possible because the program is designed to treat grammar rules as data in tables. While program modifications might be necessary for certain types of rule systems, these changes would be required only in the generation-parsing component of the system. The system's basic structure would remain constant.

The first testing of the simulation program will use, as a matter of convenience, an approximation to a stratificational model that contains dependency and phrase-structure rules and manipulates dependency networks and rules of co-occurrence to approximate relations between sememic and lexemic entities. The particular model, which I have described elsewhere,[11,12] is convenient because it is associated with an operational generation-parsing system that is ready to serve as a basic component in the simulation system.

B. UNITS OF INTERACTION

The basic units of interaction are speech forms produced in response to other speech forms. A good portion of the simulation will consist of small conversations among members of the population. A monitoring system controls the choice of interacting members.

A fundamental assumption of the simulation is that a major cause of change is the differences in the grammars of various members of a community. These differences are manifested in the varying speech forms produced during interactions. Assume that individual A has directed an utterance to individual B. B will attempt to parse the utterance with the rules available in his own recognition grammar. Each time B applies a particular rule in recognition, there might be an increase in a parameter value controlling the frequency of its usage in his generation grammar. If B's rules are not adequate for any step of the parsing, he may temporarily modify some of his own rules or temporarily borrow a rule from A in order to complete the parsing. Whether or not the temporary changes or borrowings are made permanent would be governed by other probability parameters. Changes might first be limited to the recognition grammar and permitted to enter the generation grammar only when the value of parameters sensitive to usage frequency passed a threshold. (Rules about vocabulary as well as the phonemic interpretation of phones are treated as part of the recognition- and generation-grammar systems.)

If rules pertaining to meaning are included, the conversations may be required to be coherent and to adhere to particular content areas.

C. STRUCTURE OF THE PROGRAM

The components in the system are data tables and dynamic programs.[13] One of the major data tables contains the sets of recognition and generation grammars representing the members of speech communities. Associated with each set of grammars are parameter values pertinent to the contents of the other major data table, a list of stochastic relationships applicable to a simulation. The major dynamic components are a program for parsing and generating speech forms and a monitoring system that controls the flow of the simulation. The recognition-generation component also has the task of modifying the grammars of individuals in the system. The design of this component may require alteration for simulations incorporating different theories of grammar or different notation for grammar rules belonging to the same conceptual genre. The tasks of the monitoring system include determining the passage of time and taking a periodic census to inform the experimenter of the changes that have taken place at various stages of the simulation.

III. The Modeling Process

Section II provided a description of the basic model. The term "basic" is used because the description refers to the program implementation of unalterable, primitive assumptions about the representation of members of a speech community and their mode of interaction. As indicated above, these assumptions are roughly analogous to definitions in an axiomatic system.

The analogue of axioms consists of posited stochastic relationships pertinent to the interactions among members of a community. The choice of such relationships is at the option of the researcher, and he may select them to represent a particular theory about the nature of language change and also to represent particular facts or hypotheses about historical events and social relations pertinent to a given simulation. Some typical assumptions likely to be common to many models might include:

1. A parent is more likely to speak to his child than to a member of the community selected at random.
2. A child is more likely to speak to his parent than to a member of the community selected at random.
3. A husband is more likely to speak to his wife than to a member of the community selected at random.
4. A wife is more likely to speak to her husband than to a member of the community selected at random.
5. Each time an individual interacts with a particular member of the community, the probability of future interactions with that member increases.
6. A child is more likely to adopt a grammar rule from a parent than from another member of the community selected at random.
7. An adult is less likely to adopt a grammar rule from a child than from another adult.

To incorporate the preceding assumptions in the program, the phrases "more likely" and "less likely" are redefined in terms of specific probability values, and a statement such as "the probability . . . increases" is redefined in terms of a mathematical function. Probability values are placed in the parameter lists associated with each grammar system in the community; mathematical functions that refer to the parameters are placed in the table of stochastic relationships. The number and kind of assumptions that can be incorporated in a simulation are limited only by the amount of available computer storage space, and indirectly by the availability of sufficient computer time to meet the requirements of increasingly complex simulations. For example, it is possible to model the effects of the existence of a prestige group within a community by the addition of such rules as:

8. A member of the prestige group is more likely to adopt a grammar rule from another member than from a non-member.
9. A non-member of the prestige group is more likely to adopt a grammar rule for a member than from a non-member.
10. Members of the same groups (prestige and non-prestige) are more likely to speak to each other than to members of other groups.

The experimenter may define a community subgroup by presetting pertinent parameters of the subgroup members to the same values. The treatment of multilingual contact is merely an extension of the same devices. A multilingual speaker is associated with grammars for each of his languages, and each grammar system may be associated with different parameter values. Also, special stochastic relationships may be posited for rule-borrowing between individuals speaking different languages or even for the transfer of rules between different grammar systems associated with a single individual. In general, the selection of proper parameter values and stochastic relationships should permit an experiment to model a variety of social conditions pertinent to speech interaction: marriage between speakers of different languages, sporadic interaction between members of different speech communities, even the appearance of foreign peddlers selling popular trade goods. (In this last example, the popularity of trade goods might be represented by associating a high probability of being borrowed with the names of the trade items listed in the vocabulary portion of a peddler's grammar.)

It is even possible to model the interaction of several speech communities in a particular geographical relationship. For example, consider a situation in which four speech communities, *A, B, C,* and *D*, are located so as to form the corners of a square surrounding a central community, *E.* This geographical distribution could be modeled by rules stating that interactions between members of communities *A* and *C* or *B* and *D* are less likely to occur than between members of other groups. The effects of physical barriers to communication, such as intervening rivers or mountains, could be similarly approximated.

The sudden splitting of a single speech community into two groups can be modeled by assigning zero probabilities of interaction to members of diverging groups at a specified point in time. A gradual split taking place over a lengthy period of time can be modeled by a stochastic relationship that decreases the probability of interaction as a function of elapsed time. The complementary situation in which one speech community gradually migrates into the territory of another can be modeled by the use of a function that increases the probability of interaction as a function of elapsed time.

The experimenter is also free to implement various models of individual-grammar change, for example, special hypotheses about language acquisition by children and the effects of functional load or symmetry on individual-grammar modification.

IV. Simulation Experiments

One of the major goals of this research is to perform simulations that will model language changes corresponding to events in the real world, that is, to predict a later stage of a language from a description of an earlier stage. But there are less ambitious experiments, which must be performed first, that may be of interest in themselves. For example, one must determine if the general design of the simulation system is capable of maintaining reasonable properties of language through time, both on an individual and a group basis. Conceivably, logical inconsistencies in a theoretical model, in the choice of stochastic rules, or in parameter values might cause the grammars representing the population to lose most of their rules after a few generations of interaction; or perhaps all members of the population might quickly acquire exactly the same grammars; or worse, grammars might diverge to such an extent that within a generation or two each member of the population would speak a different language.

It is also essential to determine if the simulation model can actually reflect language changes in the range of observed phenomena. For example, independent of prediction, one must determine if a model has the capability of simulating a sound shift—any sound shift, real or hypothetical.

At this stage one might check the internal validity of one's behavioral model of language-learning to insure that the development of language in the children of the simulation corresponds with language-acquisition behavior of children in the real world.

While, for a given model, there may exist combinations of parameters and rules capable of simulating acceptable real-world language change, they may be rare enough to hinder experimentation. Hopefully, this pessimistic result will not occur. I expect that preliminary experimentation with a model will yield insights about combinations of parameter values that should be avoided and about combinations that are likely to yield system behavior conforming to real-world language phenomena.

This kind of testing is much like tuning an automobile engine. The system may be extremely sensitive to particular combinations of parameter values, for example, a .5 probability of a parent interacting with his child, in combination with a .3 value of interacting with a stranger, might produce unacceptable system behavior, while any choice greater than .6 for the former and less than .2 for the latter might yield satisfactory results. In such an instance the mathematical functions pertinent to this area of interaction should be ones that do not permit the parameters to attain values outside those limits. It is likely that such a tuning will be necessary for every new modeling experiment involving different languages and/or different stochastic relationships. As part of the methodology of "tuning," one should first test the effects of only a part of the assumptions of a model, gradually adding the remainder as the more simple models are made to function satisfactorily.

Also, as indicated in Section I, it is essential to determine the effects on a simulation of different choices of random numbers. If a model is inadequate, runs differing only in the selection of random numbers may yield widely divergent behavior. The anticipated results with an adequate model would be divergent behavior—but with the divergence falling within a range too small to invalidate the model. For example, a model might be considered adequate if it predicted only hypothetical dialect variants of an attested stage of a language.

A. PREDICTION OF HISTORICAL EVENTS

One might attempt to use the simulation system to predict the future of a contemporary linguistic situation. The accuracy of the predictions would, of course, not be verifiable in the experimenter's lifetime. More fruitful experiments might involve predicting successive stages in the development of a language or language family in cases where the results could be checked against written records. Such records must be adequate for the construction of recognition and generation grammars. One would also wish to incorporate information pertaining to social structure, material culture, and geography and, if possible, detailed information about trade routes, migrations, and dated changes in social structure. If, for example, records indicate that barriers between certain social classes disappeared after a certain date, one might arrange for the program to alter the pertinent interaction parameters at the appropriate time during the course of the simulation.

In the absence of exact historical detail, one may run a simulation that posits the missing information and perhaps tests for its adequacy in accounting for future changes in a language. For example, can the simulation predict adequately if it assumes the unattested existence of trade contacts between two widely separated communities, the unattested introduction at a particular time of foreign terms for popular items of material culture, or the unattested existence of an indigenous community speaking an alien language having specific, hypothetical, but unattested grammatical features?

Ideally, results of historical-simulation studies would be adequate predictions that used only documented facts. If one is forced to incorporate speculations about history, successful prediction is not as impressive. In such cases there is justification for claiming only that the model is but one consistent, plausible theory about the factors pertinent to the language change. (It must be conceded that, at some level, a model always contains unverified speculations and that one is never justified in making a claim broader than the preceding.) If possible, one should try to predict the same results with various combinations of speculations. Each model that accurately predicts the same results is (within the limits of the simulation system) a theory about the causes of change in the test case. Analysis of runs with different models might yield information about hypotheses common to successful simulations or about the mutual incompatibility of certain combinations of hypotheses.

Another use of the program would be to test the relative validity of two hypotheses about factors of change. At best, one hypothesis would yield a valid prediction, the other fail. At worst, both would fail. More frequently, neither might yield wholly satisfactory predictions, but one prediction might be a little more accurate than the other. Note that the determination of relative accuracy might rest on many factors; for example, the only significant difference between two models might be that one predicts a verifiably false date for a minor innovation.

B. ANALYTIC SIMULATIONS

Given success in simulating historical events, one might wish to test the relative significance of various parameters in the system. Such testing, although similar to the "tuning" described in Section IV, is to be performed only after a successful predictive simulation. In essence, it would determine the range of values for a particular parameter within which the results were not significantly altered, for example, mean age at death or mean age difference between marriage partners.

Another type of simulation that must be considered analytic is the use of grammars of reconstructed languages for predicting the languages upon which the reconstructions were based. Certainly the pitfalls of circular reasoning are present for almost any conclusion to be drawn from a successful prediction. On the other hand, it is not clear to me what the significance of a failure would be. Nevertheless, assuming successful predictions have been made with real documented data, the temptation to perform such analytic experiments might be very great. Perhaps the only significance of such testing might be to determine whether the type of model necessary for successful simulation with reconstructed data were any different from that required for simulations based on attested grammars.

V. Discussion of Methodology

This paper describes a system for simulating language change within the framework of models selected at the discretion of an experimenter. Without external verification, the validity of any conclusions drawn from a simulation can be no greater than the validity of the individual assumptions incorporated in the associated model. While accurate prediction may be a criterion of success, it does not guarantee that a model accurately represents real-world events. There might exist any number of models, some mutually incompatible in their assumptions, that could yield equally accurate predictions.

Failure to predict accurately does not necessarily imply that some assumptions in a model are invalid. The model itself may have been particularly sensitive to a parameter that was not sufficiently varied in the simulations, or perhaps some highly improbable but significant event occurred in the real history of a language and was not incorporated in the set of otherwise valid assumptions of a particular model.

The ultimate function of simulation is to provide a researcher with a formal mechanism of inquiry in situations where static deductive testing of the implications of a model is not feasible because of the complexity of the phenomena involved. Explanations about historical change dependent upon unverifiable hypotheses can be tested for adequacy and internal consistency, not for validity. However, if the predictions of a simulation have been accurate, one may presume that the validity of any underlying unverifiable premises is at least as great as similar assumptions in untested models, formal or otherwise.

VI. Testing the System: Simulation of Twenty-Five Years in a Hypothetical Speech Community

It is essential to note that the simulation system and any given model of language change are separate entities. As a vehicle for testing the functioning of the simulation system, I have made use of an extremely simple model that I do not wish to defend as a real-world model of language change. Rather, its testing is to be interpreted as indicating that the simulation system works and is capable of operating with more powerful models.

A. AN ULTRA-ELEMENTARY MODEL

The initial population consisted of twenty speakers: fifteen adults and five newborn children. Age and status were the two parameters associated with each member of the community that were not directly connected with grammar rules. The age of each adult was chosen randomly. Each child was assigned age zero. The status of each adult was selected randomly.

Only phrase-structure-dependency rules were contained in the grammars. There were a total of eleven different rules contained in the community. A listing of the rules may be obtained from any of Tables 1-6. A typical rule is ART0 +*N1 N2. The existence of an equals sign between the N1 and the N2 is implied. The asterisk is data pertinent to the dependency-analysis aspect of the rule and indicates that the article is dependent on the head of the noun phrase. The dependency aspect of the rules was not pertinent to the testing of this particular model. As indicated earlier, an automatic essay-paraphrasing system that made use of dependency criteria served as the basic component for the construction of the simulation system. Although every parsing in the test runs included a dependency as well as a phrase-structure analysis, the simulation made no use of dependency criteria. The exact use of the rules in generation and parsing is described elsewhere.[11,12]

The rules governing the simulation runs included the following:

1. Probability of a speaker x speaking to an auditor y at time t:

$$\frac{1 - |(\text{status of } x \text{ at time } t) - (\text{status of } y \text{ at time } t)|}{7}$$

2. Status of speaker x at time $t + 1$ after speaking to an auditor y at time t:

$$(\text{status of } x \text{ at time } t) - \frac{(\text{status of } x \text{ at time } t) - (\text{status of } y \text{ at time } t)}{7}.$$

3. Status of auditor y at time $t + 1$ after listening to a speaker x at time t:

$$(\text{status of } y \text{ at time } t) - \frac{(\text{status of } y \text{ at time } t) - (\text{status of } x \text{ at time } t)}{4}$$

4. Status, at time $t + 1$, of potential participants in a conversation at time t who did not converse: + 0.01 for the individual of greater status; — 0.01 for the individual of lesser status.

5. Status of a newborn child: a random value between 0.01 and 0.99.

6. Frequency weight of a grammar rule m at time $t + 1$ that was used one or more times in the parsing of a single sentence at time t:

$$(\text{frequency weight of } m \text{ at time } t) + 0.03\ x\ (\text{subscript of the right half of rule}).$$

The computation is applied repeatedly during time interval t for as many sentences as there are in the discourse.

7. Frequency weight of a grammar rule m at time $t + 1$ that was not used in the parsing of a single sentence during time interval t:

$$(\text{frequency of } m \text{ at time } t) - (\text{an average decrement of } 0.003);$$

that is, there is a 30 per cent chance of a 0.01 decrement. The computation is applied repeatedly during time interval t for as many sentences as there are in the discourse not pertinent to rule m.

8. Threshold frequency weight for adding or removing a rule from a grammar: 0.02.

9. Initial frequency weight of a rule borrowed by an individual under two years of age: 0.20; over two years of age: 0.40.

10. Probability of death for an individual in a given year: age/1,000 for speakers over ten years of age, 0.10 for speakers ten years and under.

Except in the case of rule 4, all computed values greater than 0.99 are rounded to 0.99; values computed as less than 0.01 are rounded to 0.01. In the case of rule 4, the rounding is to 0.98 and 0.02, respectively. Also, no distinction between generation and recognition grammars was made with reference to the status of rules; a rule was either in a particular grammar for both generation and parsing or not present at all.

The flow of the group interaction can be described in terms of major and minor cycles. Each member of the population is assigned a number. A major cycle is begun by picking the first member as speaker. The second member of the population is then considered a potential auditor. Whether or not he is selected is determined by the first rule and reference to a random-number generator. Whether or not a conversation takes place, the clock of the system is incremented by one minimal time unit. The process is repeated for the third and successive members of the community. When each member of the community has been considered as a potential auditor of the speaker, a minor cycle has been completed. The second member of the population is then selected as speaker of the next minor cycle. When every member of the community has served as speaker for a minor cycle, a major cycle has been completed. One major cycle is equivalent to one year. The number of minimal time units in a minor cycle is equal to the number of individuals in the population—in this case, twenty.

The birth rate in the model is identical to the death rate. The probability of death for an individual is computed each time he is selected as speaker for a minor cycle. If a random number falls within the appropriate range, that individual dies before he has a chance to

TABLE 1

Census: Year 1

Rule			Mean Frequency Weight in Total Population	Speakers	Mean Frequency Weight among Speakers
Run 1:					
ART0	+ * N1	N2	00064	00014	00130
ADJ0	+ * N1	N1	00100	00017	00125
N2	* + SUBCN1	N3	00021	00004	00127
N2	* + MOD1	N3	00037	00007	00133
	N0	N1	00004	00024	00004
V2	* + N4	V3	00074	00015	00134
	V0	V1	00004	00024	00004
PART0	* + N4	MOD1	00025	00005	00124
PREP0	* + N4	MOD1	00031	00006	00123
N4	* + V3	S1	00004	00024	00004
PNRCN0	* + N3	SUBCN1	00021	00004	00125
Run 2:					
ART0	+ * N1	N2	00064	00014	00130
ADJ0	+ * N1	N1	00100	00017	00125
N2	* + SUBCN1	N3	00021	00004	00127
N2	* + MOD1	N3	00037	00007	00133
	N0	N1	00004	00024	00004
V2	* + N4	V3	0007[illegible]	00015	00134
	V0	V1	00004	00024	00004
PART0	* + N4	MOD1	00025	00005	00124
PREP0	* + N4	MOD1	00031	00006	00123
N4	* + V3	S1	00004	00024	00004
PNRCN0	* + N3	SUBCN1	00021	00004	00125
Run 3:					
ART0	+ * N1	N2	00064	00014	00130
ADJ0	+ * N1	N1	00100	00017	00125
N2	* + SUBCN1	N3	00021	00004	00127
N2	* + MOD1	N3	00037	00007	00133
	N0	N1	00004	00024	00004
V2	* + N4	V3	00074	00015	00134
	V0	V1	00004	00024	00004
PART0	* + N4	MOD1	00025	00005	00124
PREP0	* + N4	MOD1	00031	00006	00123
N4	* + V3	S1	00004	00024	00004
PNRCN0	* + N3	SUBCN1	00021	00004	00125

talk. He is immediately replaced by a newborn child with the same number, an age of zero, and a randomly determined status.

Newborn children in this particular model do not have completely empty grammars. Rather, they are assigned that minimum of rules to generate the simplest well-formed sentence: N4* + V3 = S1, N0 = N1, and V0 = V1. Their inclusion does not indicate the author's commitment to any theory of innate ideas but rather was necessary as a programing expedient. The frequency weight permanently assigned to these rules was 0.04.

B. TESTING THE MODEL

The exact forms of the rules of the model, especially the values of constants, were selected after much trial and error. The goal of the testing was to attain a situation of stability for the mean frequency weights of the grammar rules. Early versions of the model rules led to

TABLE 2

Census: Year 6

Rule				Mean Frequency Weight in Total Population	Speakers	Mean Frequency Weight among Speakers
Run 1:						
ARTO	+ *	N1	N2	00117	00024	00117
ADJ0	+ *	N1	N1	00112	00024	00112
N2	* +	SUBCN1	N3	00061	00021	00072
N2	* +	MOD1	N3	00103	00021	00117
		N0	N1	00004	00024	00004
V2	* +	N4	V3	00127	00024	00127
		V0	V1	00004	00024	00004
PART0	* +	N4	MOD1	00043	00015	00066
PREP0	* +	N4	MOD1	00054	00020	00067
N4	* +	V3	S1	00004	00024	00004
PNRCN0	* +	N3	SUBCN1	00052	00021	00062
Run 2:						
ARTO	+ *	N1	N2	00121	00023	00125
ADJ0	+ *	N1	N1	00114	00023	00120
N2	* +	SUBCN1	N3	00055	00017	00074
N2	* +	MOD1	N3	00120	00022	00131
		N0	N1	00004	00024	00004
V2	* +	N4	V3	00127	00023	00133
		V0	V1	00004	00024	00004
PART0	* +	N4	MOD1	00064	00022	00072
PREP0	* +	N4	MOD1	00061	00021	00071
N4	* +	V3	S1	00004	00024	00004
PNRCN0	* +	N3	SUBCN1	00045	00017	00062
Run 3:						
ARTO	+ *	N1	N2	00134	00024	00134
ADJ0	+ *	N1	N1	00125	00024	00125
N2	* +	SUBCN1	N3	00050	00017	00066
N2	* +	MOD1	N3	00133	00024	00133
		N0	N1	00004	00024	00004
V2	* +	N4	V3	00140	00024	00140
		V0	V1	00004	00024	00004
PART0	* +	N4	MOD1	00063	00024	00063
PREP0	* +	N4	MOD1	00065	00024	00065
N4	* +	V3	S1	00004	00024	00004
PNRCN0	* +	N3	SUBCN1	00043	00017	00056

loss of all grammar rules, to attainment of maximum frequency weight for every rule, or to some combination of factors that led to maximization of frequency weights for some grammar rules and loss for others.

The current model is of such a nature that the frequency weights of most grammar rules would reach asymptotes of 0.99 were it not for the fact that the death rate is such that individuals usually die before the weights of their rules all reach such values.

Tables 1-6 contain results of censuses taken every five years during a span of twenty-five years for each of three separate runs. Each census indicates the number of speakers possessing each grammar rule, the mean frequency of each rule among speakers actually possessing it, and the mean frequency of each rule in the total population. The censuses in the tables were constructed from actual computer output, and all values are expressed as octal integers. To convert such values to the decimal system, multiply each integer going from right to left by successive powers of eight,

TABLE 3

CENSUS: YEAR 11

Rule			Mean Frequency Weight in Total Population	Speakers	Mean Frequency Weight among Speakers
Run 1:					
ARTO	+ * N1	N2	00125	00023	00132
ADJ0	+ * N1	N1	00116	00024	00116
N2	* + SUBCN1	N3	00100	00022	00107
N2	* + MOD1	N3	00124	00022	00136
	N0	N1	00004	00024	00004
V2	* + N4	V3	00134	00024	00134
	V0	V1	00004	00024	00004
PARTO	* + N4	MOD1	00062	00022	00070
PREP0	* + N4	MOD1	00057	00022	00065
N4	* + V3	S1	00004	00024	00004
PNRCN0	* + N3	SUBCN1	00056	00022	00063
Run 2:					
ARTO	+ * N1	N2	00127	00023	00134
ADJ0	+ * N1	N1	00115	00023	00122
N2	* + SUBCN1	N3	00067	00022	00075
N2	* + MOD1	N3	00130	00023	00134
	N0	N1	00004	00024	00004
V2	* + N4	V3	00133	00023	00140
	V0	V1	00004	00024	00004
PARTO	* + N4	MOD1	00067	00022	00075
PREP0	* + N4	MOD1	00062	00023	00065
N4	* + V3	S1	00004	00024	00004
PNRCN0	* + N3	SUBCN1	00046	00022	00053
Run 3:					
ARTO	+ * N1	N2	00123	00023	00130
ADJ0	+ * N1	N1	00116	00023	00123
N2	* + SUBCN1	N3	00055	00017	00074
N2	* + MOD1	N3	00126	00023	00133
	N0	N1	00004	00024	00004
V2	* + N4	V3	00133	00023	00137
	V0	V1	00004	00024	00004
PARTO	* + N4	MOD1	00060	00023	00063
PREP0	* + N4	MOD1	00060	00023	00063
N4	* + V3	S1	00004	00024	00004
PNRCN0	* + N3	SUBCN1	00041	00017	00054

for example, an octal integer, 132, may be converted to the decimal system as follows: $2 \times 8^0 + 3 \times 8^1 + 1 \times 8^2 = 2 \times 1 + 3 \times 8 + 1 \times 64 = 90$ in the decimal system. The number of speakers indicated in the censuses is always an integer. The frequency weights, although expressed as integers, are to be treated as decimal fractions in the range 0.01-0.99 after the conversion from octal integer to decimal integer has been completed. Thus, a value of 143 in a census table is to be ultimately interpreted as the decimal value, 0.99.

Figures 1—4 contain graphs of the mean frequencies in the total population for selected rules (on the basis of yearly censuses). Figures 5-8 contain graphs representing the number of speakers possessing the rules mentioned in Figures 1-4 (also on the basis of yearly censuses).

The frequency increment of a rule used in paraphrasing is, as rule 6 of the model rules indicates, a function of the subscript of the right half of a grammar rule. The subscripts control the order of applica-

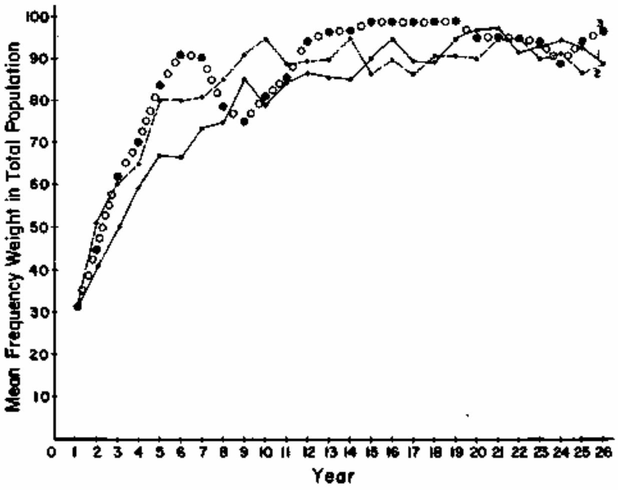

FIG. 1.—Mean frequencies in the total population for the rule N2 * + MOD1 = N3 on the basis of yearly censuses. Three runs.

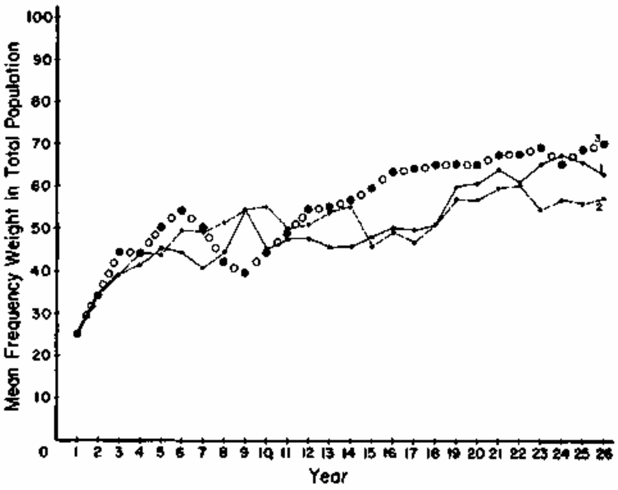

FIG. 2.—Mean frequencies in the total population for the rule PREP0 * + N4 = MOD1 on the basis of yearly censuses. Three runs.

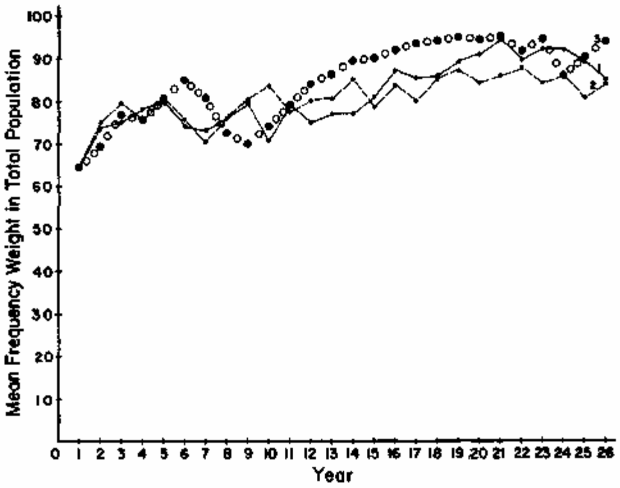

FIG. 3.—Mean frequencies in the total population for the rule ADJ0 + * N1 = N1 on the basis of yearly censuses. Three runs.

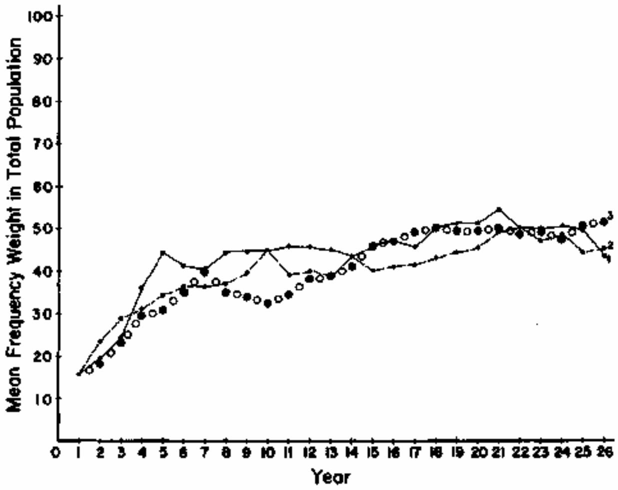

Fig. 4.—Mean frequencies in the total population for the rule PNRCN0 * + N3 = SUBCN1 on the basis of yearly censuses. Three runs.

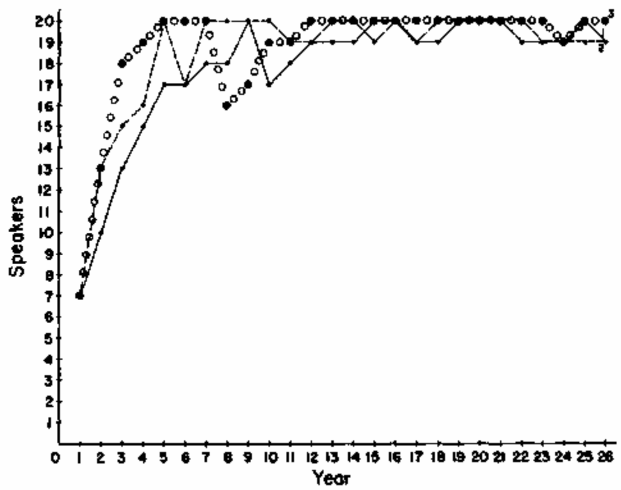

Fig. 5.—The number of speakers possessing the rule N2 * + MOD1 = N3 on the basis of yearly censuses. Three runs.

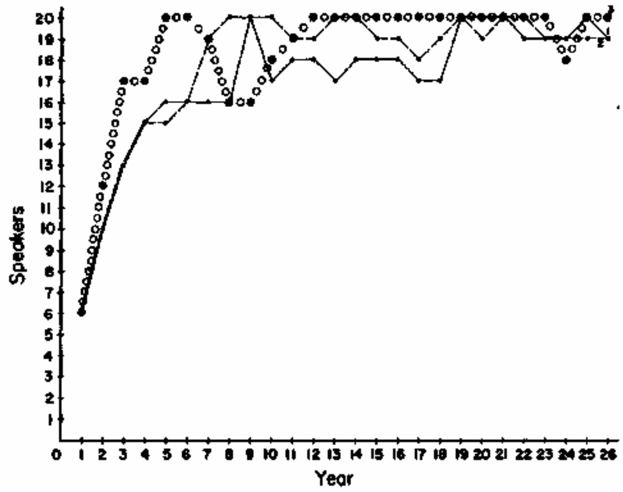

Fig. 6.—The number of speakers possessing the rule PREP0 * + N4 = MOD1 on the basis of yearly censuses. Three runs.

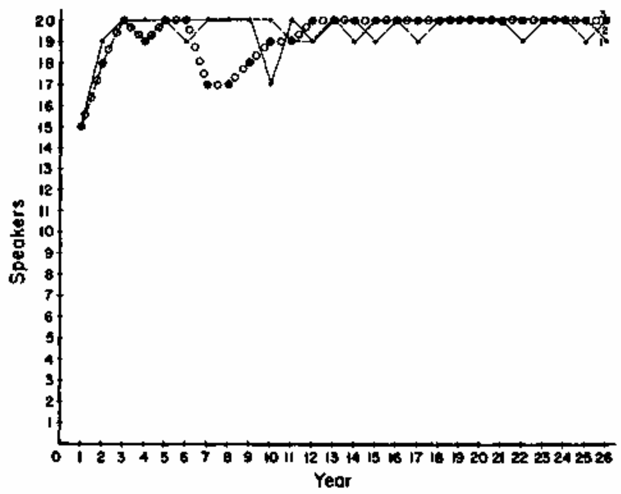

Fig. 7.—The number of speakers possessing the rule ADJ0 + * N1 = N1 on the basis of yearly censuses. Three runs.

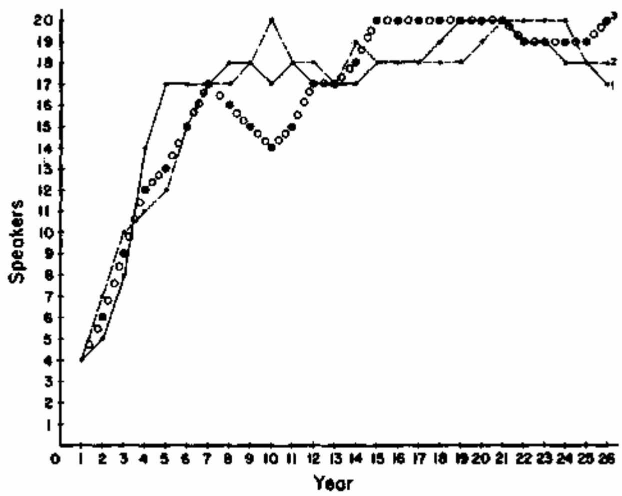

Fig. 8.—The number of speakers possessing the rule PNRCN0 * + N3 = SUBCN1 on the basis of yearly censuses. Three runs.

TABLE 4

CENSUS: YEAR 16

Rule			Mean Frequency Weight in Total Population	Speakers	Mean Frequency Weight among Speakers
Run 1:					
ART0	+ * N1	N2	00135	00024	00135
ADJ0	+ * N1	N1	00127	00024	00127
N2	* + SUBCN1	N3	00105	00022	00115
N2	* + MOD1	N3	00136	00024	00136
	N0	N1	00004	00024	00004
V2	* + N4	V3	00137	00024	00137
	V0	V1	00004	00024	00004
PART0	* + N4	MOD1	00072	00024	00072
PREP0	* + N4	MOD1	00062	00022	00067
N4	* + V3	S1	00004	00024	00004
PNRCN0	* + N3	SUBCN1	00057	00022	00065
Run 2:					
ART0	+ * N1	N2	00131	00024	00131
ADJ0	+ * N1	N1	00123	00024	00123
N2	* + SUBCN1	N3	00104	00022	00114
N2	* + MOD1	N3	00131	00024	00131
	N0	N1	00004	00024	00004
V2	* + N4	V3	00135	00024	00135
	V0	V1	00004	00024	00004
PART0	* + N4	MOD1	00065	00023	00070
PREP0	* + N4	MOD1	00061	00023	00063
N4	* + V3	S1	00004	00024	00004
PNRCN0	* + N3	SUBCN1	00052	00022	00056
Run 3:					
ART0	+ * N1	N2	00142	00024	00142
ADJ0	+ * N1	N1	00134	00024	00134
N2	* + SUBCN1	N3	00107	00024	00107
N2	* + MOD1	N3	00142	00024	00142
	N0	N1	00004	00024	00004
V2	* + N4	V3	00143	00024	00143
	V0	V1	00004	00024	00004
PART0	* + N4	MOD1	00075	00024	00075
PREP0	* + N4	MOD1	00075	00024	00075
N4	* + V3	S1	00004	00024	00004
PNRCN0	* + N3	SUBCN1	00057	00024	00057

tion of the rules in parsing and generation. The use of subscripts as a factor in computing frequency-weight increment was an empirical attempt to reflect the tendency of some high subscript rules to have a much lower frequency weight than those with lesser subscripts. The decrement for weights of rules not used in parsing does not involve subscripts. It was necessary to keep the frequency weight of the terminal rules, N0 = N1 and V0 = V1, a low constant value to prevent the loss of most other rules from each grammar.

As indicated, a total of three runs was performed with the model. They differed only in the choice of random numbers presented to the decision-making portions of the program. The initial populations in each run were identical in composition. The creation of the starting population was accomplished as follows:

An additional speaker, possessing every rule in the system (with randomly assigned frequency weights) was set to converse with every other individual in the population in a preprocessing minor cycle. (Newborn

TABLE 5

CENSUS: YEAR 21

Rule				Mean Frequency Weight in Total Population	Speakers	Mean Frequency Weight among Speakers
Run 1:						
ART0	+ *	N1	N2	00137	00024	00137
ADJ0	+ *	N1	N1	00135	00024	00135
N2	* +	SUBCN1	N3	00122	00024	00122
N2	* +	MOD1	N3	00141	00024	00141
		N0	N1	00004	00024	00004
V2	* +	N4	V3	00140	00024	00140
		V0	V1	00004	00024	00004
PART0	* +	N4	MOD1	00100	00024	00100
PREP0	* +	N4	MOD1	00077	00024	00077
N4	* +	V3	S1	00004	00024	00004
PNRCN0	* +	N3	SUBCN1	00065	00024	00065
Run 2:						
ART0	+ *	N1	N2	00133	00024	00133
ADJ0	+ *	N1	N1	00127	00024	00127
N2	* +	SUBCN1	N3	00121	00024	00121
N2	* +	MOD1	N3	00135	00024	00135
		N0	N1	00004	00024	00004
V2	* +	N4	V3	00140	00024	00140
		V0	V1	00004	00024	00004
PART0	* +	N4	MOD1	00075	00024	00075
PREP0	* +	N4	MOD1	00073	00024	00073
N4	* +	V3	S1	00004	00024	00004
PNRCN0	* +	N3	SUBCN1	00060	00024	00060
Run 3:						
ART0	+ *	N1	N2	00137	00024	00137
ADJ0	+ *	N1	N1	00136	00024	00136
N2	* +	SUBCN1	N3	00116	00024	00116
N2	* +	MOD1	N3	00136	00024	00136
		N0	N1	00004	00024	00004
V2	* +	N4	V3	00140	00024	00140
		V0	V1	00004	00024	00004
PART0	* +	N4	MOD1	00077	00023	00103
PREP0	* +	N4	MOD1	00103	00024	00103
N4	* +	V3	S1	00004	00024	00004
PNRCN0	* +	N3	SUBCN1	00061	00024	00061

babies were omitted.) Rule 1 of the model, governing probability of interaction, did not apply. Each auditor had only the grammar rules of a newborn child; a preassigned, randomly determined status; and a randomly determined age. Rules borrowed by auditors entered their grammars with a frequency weight equal to 65 plus a randomly determined value between 0 and 30. After the initializing minor cycle, the primordal speaker was eliminated from the system. I assume no responsibility for the philosophical implications of this method of creating a starting population.

The initializing procedure was identical for each of the three trial runs, which were permitted to deviate from one another subsequently. While the fate of various individuals differed widely in each run, the mean frequencies computed in the censuses appear quite close at identical time periods. What of course is meant by "close"? Statistical interpretation of the results is complicated by the problem of choosing a pertinent test. Should the population in the various runs be

TABLE 6

CENSUS: YEAR 26

Rule			Mean Frequency Weight in Total Population	Speakers	Mean Frequency Weight among Speakers
Run 1:					
ART0	+ * N1	N2	00130	00023	00134
ADJ0	+ * N1	N1	00125	00023	00132
N2	* + SUBCN1	N3	00111	00021	00126
N2	* + MOD1	N3	00130	00023	00135
	N0	N1	00004	00024	00004
V2	* + N4	V3	00132	00023	00137
	V0	V1	00004	00024	00004
PART0	* + N4	MOD1	00077	00023	00103
PREP0	* + N4	MOD1	00076	00023	00101
N4	* + V3	S1	00004	00024	00004
PNRCN0	* + N3	SUBCN1	00053	00021	00063
Run 2:					
ART0	+ * N1	N2	00131	00024	00131
ADJ0	+ * N1	N1	00124	00024	00124
N2	* + SUBCN1	N3	00114	00022	00124
N2	* + MOD1	N3	00130	00023	00135
	N0	N1	00004	00024	00004
V2	* + N4	V3	00134	00024	00134
	V0	V1	00004	00024	00004
PART0	* + N4	MOD1	00075	00023	00100
PREP0	* + N4	MOD1	00071	00023	00074
N4	* + V3	S1	00004	00024	00004
PNRCN0	* + N3	SUBCN1	00055	00022	00062
Run 3:					
ART0	+ * N1	N2	00140	00024	00140
ADJ0	+ * N1	N1	00135	00024	00135
N2	* + SUBCN1	N3	00124	00024	00124
N2	* + MOD1	N3	00141	00024	00141
	N0	N1	00004	00024	00004
V2	* + N4	V3	00143	00024	00143
	V0	V1	00004	00024	00004
PART0	* + N4	MOD1	00105	00024	00105
PREP0	* + N4	MOD1	00106	00024	00106
N4	* + V3	S1	00004	00024	00004
PNRCN0	* + N3	SUBCN1	00063	00024	00063

treated as a sample of the total group? If a sample, from what size population? One might compute the mean frequency of all census mean frequency values at each time interval, then check to see if any individual values fall outside a computed standard error. But this is a weak test. Its use might indicate success where a linguist might judge failure; for example, a linguist might feel the linguistic situations emergent from different trials were too divergent to be considered as variants of the same language, even though all census values fell within the range of the standard error. A graphic display of the results may present evidence as least as convincing as any statistical test. In any case, a sample of three runs is too small for any statistical test to be of much significance. In my opinion, the graphs in Figures 1-8 are sufficiently convincing that the claim for similar results at similar time intervals is justified. The graphs also suggest that a near equilibrium state was attained in the later years of each run. The sharp rise in mean frequency weights at the beginning of each run is most likely due to the random and independent assignment of frequency weights to individual rules. The rules in a grammar are not independent of each other. Neither is their usage in a generation system. Accordingly, the initial conditions were unstable. The functioning of the system seemed to force the values onto stable levels.

As indicated earlier, the purpose of the computer test was to check out the simulation system rather than the model. I believe that the results are a positive indication of the feasibility of simulating group language behavior within the conceptual framework described in Sections I-V of this paper. In other words, while the model is trivial, the simulation system and the implied methodology are not.

Received May 23, 1966

References

1. Orcutt, B., Greenberger, M., Korbel, J., and Rivlin, A. *Microanalysis of Socioeconomic Systems: A Simulation Study*. New York: Harper & Row, 1961.
2. Bush, R., and Mosteller, F. *Stochastic Models of Learning*. New York: John Wiley & Sons, 1955.
3. Guetzkow, H. (ed.). *Simulation in Social Science: Readings*. Englewood Cliffs, N.J.: Prentice-Hall, Inc., 1962.
4. Hoggat, A. C., and Balderston, F. E. (eds.). *Symposium on Simulation Models: Methodology and Applications in the Behavioral Sciences*. Cincinnati: Southwestern Publishing Co., 1963.
5. Feigenbaum, E. A., and Feldman, J. (eds.). *Computers and Thought*. New York: McGraw-Hill Book Co., 1963.
6. Hymes, Dell (ed.). *Uses of Computers in Anthropology*. The Hague: Mouton & Co., 1965.
7. McPhee, W. N. *Formal Theories of Mass Communication*. New York: Free Press, 1963.
8. Gilbert, J. P., and Hammel, E. A., Jr. "Computer Simulation and Analysis of Problems in Kinship and Social Structure," *American Anthropologist,* Vol. 68, No. 1 (February, 1966).
9. Sapir, E. *Language*. New York: Harcourt, Brace & World, 1921.
10. Bloomfield, L. *Language*. New York: Holt, Rinehart & Winston, 1933.
11. Klein, S. "Automatic Paraphrasing in Essay Format," *Mechanical Translation,* Vol. 8, Nos. 3 and 4 (June and October, 1965).
12. ———. "Control of Style with a Generative Grammar," *Language,* Vol. 41, No. 4 (October-December, 1965).
13. ——— "Some Components of a Program for Dynamic Modelling of Historical Change in Language." (Preprints of Invited Papers for 1965, Paper No. 14.) International Conference on Computational Linguistics, New York, May 19-21, 1965.