mone Research, S. Andersson, Ed. (Almqvist & Wiksell, Stockholm, 1973), pp. 137–142.
9. R. S. Yalow and S. A. Berson, in *Structure-Activity Relationships of Protein and Polypeptide Hormones*, M. Margoulies and F. C. Greenwood, Eds. (Excerpta Medica, Amsterdam, 1971), part 1, pp. 48–56.
10. W. A. Russel-Hunter, *Biology of Lower Vertebrates* (Macmillan, New York, 1968), p. 146; C. L. Prosser and F. A. Brown, *Comparative Animal Physiology* (Saunders, Philadelphia, 1965), p. 112.
11. This work was supported by VA project 9678-01 and Mt. Sinai Clinical Genetics Center grant 19-443.

28 July 1975; revised 25 August 1975

Vocabulary Richness: A Sociolinguistic Analysis

Abstract. *A regression analysis of lexical diversity in the informal speech of 120 mature speakers of French in Montreal reveals no direct effect of socioeconomic level or residential milieu. All social effects are mediated by a single variable: educational attainment. The analysis also confirms a continuing enrichment of productive vocabulary with increasing age.*

The linguistic correlates of social stratification have provoked a variety of evaluative attitudes about the nature of "upper-class" versus "lower-class" speech. These range from the verbal deprivation (*1*) and distorted communication (*2*) descriptions of ghetto or working-class speakers, through the more subtle assessments of differentiated functional capacities of "elaborated and restricted codes" (*3*), and finally to the relativism characterizing much urban sociolinguistics (*4*). In our work on a million-word corpus of spoken French in Montreal (*5*), focusing on phonological (*6*), syntactic (*7*), and semantic (*8*) differentiation according to socioeconomic and demographic characteristics of 120 speakers, we have found no justification for positively or negatively evaluating any of the linguistic attributes that distinguish among these speakers. We have begun a computer-assisted analysis of lexical usage in the corpus, and in this area there do exist quantifiable notions such as the range or richness of vocabulary which might provide grounds, however superficial, for such evaluation. We report here the first results of this study.

The quantitative data we analyze are the number of different words, D, in each of the 120 interviews, compared to the total number of words, T, in the same interviews, and six socioeconomic and demographic factors pertaining to each speaker. The interviews averaged about 1 hour and all covered the same broad topic, everyday life in Montreal, past and present, a subject with which all speakers were equally familiar. The word counts, from a computer-stored transcription, were not corrected for homonymy, nor for differently conjugated or declined forms of the same root, but these sources of error apply to all of the speakers in a presumably uniform way.

In general, the number of different words in a text is a nonlinear function of its total length. In our data it appears to be of the form $D = \mathrm{a} + \mathrm{b}T^{\mathrm{r}}$ where a and b are constants, and r is approximately 0.7. We have repeated the analyses reported below using a range of values of r between 0.65 and 0.75, without substantially changing the results.

The demographic data include age (from 15 to 85 years) and sex, and the socioeconomic factors are mean income of speaker's residential area, educational level attained, and index of occupational status of the head of speaker's household, and a similar index for the speaker himself. Needless to say, these last four are highly correlated. To find which factors contribute independently to D in a significant way, we carried out a multiple regression of D on T^{r} and the six variables, allowing for quadratic terms and products, and using the "forward selection procedure" (*9*) to determine which coefficients are statistically significant.

Apart from T^{r}, only one term has a significant coefficient ($P < .05$), namely the product of age, A, and educational level, E. Even when the other factors are "forced" into the regression, they have small coefficients, while the age-education interaction effect remains large and significant. Detailed investigation of this interaction results in the regression

$$D = 24.7 + 1.912T^{0.7} + 2.775A^{0.7}E$$

where A is measured in years and E in coding categories 1 to 5, each representing an increment of about 4 years' schooling. This formula indicates that in the population represented by our sample, each person incorporates new words into his productive vocabulary at a slowly decreasing rate over time, but this rate can be magnified up to five times through extensive education.

Allowing less strict significance criteria ($P < .1$) does not enable any of the factors other than A or E to attain significance. By allowing higher order terms in A and E, such as A^2E, we can modify the regression function somewhat so that D initially rises faster in the case of highly educated speakers, peaks near age 50, and then declines slightly, but this effect is of borderline significance. Within each of the five educational categories considered separately, the data contain too much scatter to confirm the relationship between A and D in any detail. However, for all five categories, by forcing A into the regression we always obtain a positive coefficient and preclude the significance of any of the other social or demographic factors.

The important fact that emerges from this analysis is that any contribution to richness of vocabulary (as indicated by D) from such socioeconomic factors as residential milieu or occupational status of parent are completely accounted for by the effects of these factors on educational attainment. This lends no support to theories that linguistic competence is degraded by an "impoverished" childhood environment.

Also of interest is the importance of age. Enrichment of productive vocabulary continues at least until age 50.

These two observations tie in with other work on the same speech community. Linguistic competence with respect to syntax of children from different socioeconomic milieus (but in the same school grade) shows little systematic variation (*10*). The lexicon is among the most malleable components of a language and the most responsive to the circumstances of language use (*11*). This is illustrated by the systematic, although highly variable acquisition rates we find among adults. Such acquisition does not generally occur among adults, as far as is known, with respect to phonology or syntax. This malleability of the lexicon explains why the language of educated people can be lexically richer, while no comparable objective measure exists which might reveal a "better" phonology or a more functional or logical syntax.

DAVID SANKOFF
RÉJEAN LESSARD
Centre de Recherches Mathématiques, Université de Montréal, C.P. 6128, Montréal H3C 3J7, Québec, Canada

References and Notes

1. C. Bereiter, S. Engelmann, J. Osborn, P. A. Reidford, in *Pre-School Education Today*, F. M. Hechinger, Ed. (Doubleday, New York, 1966), pp. 105–135.
2. C. Mueller, in *Recent Sociology, No. 2: Patterns of Communicative Behavior*, H. P. Dreitzel, Ed. (Macmillan, New York, 1970), pp. 101–113.
3. B. Bernstein, *Class, Codes and Control* (Routledge, London, 1971), vols. 1 and 2.
4. W. Labov, *Language in the Inner City* (Univ. of Pennsylvania Press, Philadelphia, 1973); P. Trudgill, *The Social Differentiation of English in Norwich* (Cambridge Univ. Press, London, 1974); H. J. Cedergren, thesis, Cornell University (1973).
5. G. Sankoff and H. J. Cedergren, *Lang. Soc.* **1**, 173 (1972); D. Sankoff and G. Sankoff, in *Canadian Languages in Their Social Context*, R. Darnell, Ed. (Linguistic Research, Edmonton, 1973), pp. 7–64.
6. G. Sankoff and H. J. Cedergren, in *Linguistic Diversity in Canadian Society*, R. Darnell, Ed. (Linguistic Research, Edmonton, 1971), pp. 61–87.
7. G. Sankoff, in *New Ways of Analyzing Variation*

in English, C.-J. N. Bailey and R. W. Shuy, Eds. (Georgetown Univ. Press, Washington, D.C., 1973), pp. 44–66.
8. S. Laberge, thesis, Université de Montréal (1972).
9. N. R. Draper and H. Smith, *Applied Regression Analysis* (Wiley, New York, 1966), p. 169.
10. M. Pagé, thesis, Université de Montréal (1971).
11. Linguistic theory has even entertained the possibility, in certain contact situations, of the relexification of a language, that is, the replacement of the bulk of its vocabulary with words from another language while retaining much of the original phonology, syntax, and semantics.
12. Supported in part by Canada Council grant S72-0988 and National Research Council of Canada grant A8867. We thank P. Bratley and F. Ouellette for the use of their text-handling system JEUDEMO.

20 March 1975; revised 2 June 1975

Auto-Shaping in Rats to the Presentation of Another Rat Predicting Food

Abstract. *Rats direct social rather than eating behavior toward a stimulus rat that predicts the imminent delivery of food. This result suggests that a predictive stimulus does not become a substitute for a reward, but its characteristics elicit and support a particular subset of the responses commonly related to that reward.*

Brown and Jenkins (*1*) discovered that pigeons exposed to repeated pairings of a lighted key with food began to peck the lighted key. They termed this phenomenon auto-shaping because the pigeons pecked even though food delivery was always independent of their behavior. Hearst and Jenkins (*2*) summarized subsequent studies showing that auto-shaping is a robust phenomenon which occurs with rewards other than food and in species other than pigeons.

The parallels between auto-shaping and classical conditioning led many investigators to analyze examples of auto-shaping from the viewpoint of stimulus substitution, an explanation proposed by Pavlov for classical conditioning of the salivary reflex in dogs (*2–4*). According to the stimulus substitution hypothesis, the subject responds to the predictive stimulus (CS) as if it were a substitute or surrogate for the reward (US). Support for the stimulus substitution explanation of auto-shaping was provided by experiments showing that the topography of behavior directed toward the predictive stimulus closely resembles that of behavior elicited by the reward. For example, male pigeons direct brief, hard pecks at a key which signals food, soft "sipping" pecks at a key which signals water, and court a key which predicts access to a receptive female pigeon (*2, 4*). Rats bite and lick a bar which predicts food (*3*).

However, in other experiments behavior elicited by the predictive stimulus did not resemble behavior elicited by the reward (*2–5*). Most recently, Wasserman (*5*) found that baby chicks in a refrigerated chamber approached, pecked, and "snuggled" at a lighted key which predicted the onset of heat, although the chicks' behavior in the presence of heat consisted of immobility and wing extension. Hogan (*6*) accounted for the chicks' behavior by noting that in a cool environment chicks approach, peck, and "snuggle" at a mother hen, thereby inducing her to brood them.

Hogan's explanation suggests a systematic exception to the stimulus substitution hypothesis. Behavior elicited by a predictive stimulus will resemble behavior elicited by the reward only to the extent that the predictive stimulus is compatible with and supports such behavior. If the predictive stimulus elicits and supports other behavior commonly related to the reward, conditioning will occur, but the behavior elicited by the predictive stimulus will differ from that occurring in response to the reward. From this viewpoint, Wasserman's results contradicted the stimulus substitution hypothesis, because the key light, set in a wall, most adequately supported those aspects of thermoregulatory behavior in chicks commonly directed toward the mother hen.

Table 1. Median number of trials on which subjects showed different social contact behaviors over the last 5 days of acquisition (days 7 to 11). The maximum score on any day was 30. Only subjects in groups CS^+ and CS^s are included because, with the exception of one animal, the median scores of subjects in the CS^r and CS^w groups were zero for all behaviors.

Subject	Social contact behaviors			
	Paw	Groom	Crawl-over	Ano-genital sniff
		Group: CS^+		
1	25	18	21	0
2	23	20	3	4
3	26	21	10	1
4	19	16	4	0
5	20	14	6	5
Mean	*22.6*	*17.8*	*8.8*	*2.0*
		Group: CS^s		
1	0	0	0	0
2	4	0	0	0
3	0	0	0	0
4	14	8	1	2
5	6	2	0	0
Mean	*4.8*	*2.0*	*0.2*	*0.4*

In the present experiment we examined another case in which the predictive stimulus and the reward supported different, although potentially related, behaviors. We employed a live rat fastened to a platform as a predictive stimulus for food. If the stimulus substitution hypothesis is correct, subject rats should treat the predictive rat as food, gnawing or biting it. However, social behavior elicited and supported by the predictive rat is also potentially related to feeding in the rat. Rats feed together, follow each other to food, and, both as pups and adults, learn feeding locations by approaching other rats (*7*). On this basis, the subjects would be expected to incorporate the predictive rat into a social feeding pattern, increasing their frequency of approach and social contact.

To test these hypotheses we compared the behavior of an experimental group with that of three control groups. Each group consisted of five male Wistar albino rats, 90 days of age. During acquisition each rat received 30 10-second presentations of the predictive stimulus (CS) on a variable time schedule with a mean interstimulus interval of 60 seconds (VT 60 seconds). The stimulus platform, driven by a synchronous motor and cam assembly, was presented sideways through a flap door adjacent to the wall containing the food tray. For the experimental group (CS^+ group) each presentation of the predictive rat was followed by one 45-mg food pellet. The CS^s (social) group received the same pattern of presentation of the stimulus rat, but no food was ever delivered. Since rats are highly social (*8*), this group served as a baseline of social reactivity to the stimulus rat. The CS^r (random) group was presented with the stimulus rat and food randomly on two independent variable-time 60-second programs. The purpose of this group was to determine the importance of the pairings of the stimulus rat and food (*9*). The CS^w (wood) group was subject to the same procedures as the CS^+ group, except that the predictive stimulus was a rat-sized block of wood fastened to the platform. The purpose of this group was to separate the social and predictive effects of the stimulus rat. Rats in the CS^+ group might approach the stimulus rat because of its predictive quality and then engage in social contact because of their proximity to the stimulus rat.

All rats were housed alone during the experiment. After adaptation to a 23-hour feeding schedule, each rat received 22 days of training: 3 days of pretraining, 11 days of acquisition, and 8 days of extinction. On the first day of pretraining we exposed each subject to the experimental chamber for 30 minutes; on the next day we trained