The effects of anxiety and time upon productivity and vocabulary diversity in the absence of nonlexical feedback

JACK N. HANEY, Ohio University, Athens, Ohio 45701

Feedback variables normally associated with the interview situation were controlled by having Ss write their responses as a function of high and low anxiety-arousing instructions and in 2- or 10-min periods. Productivity was found to be significantly inhibited for those groups operating under anxiety-arousing conditions, while vocabulary diversity was restricted during short response periods. A within-S analysis of type-token ratios revealed a significant and systematic increase in this variable over time in the absence of interpersonal feedback.

Lennard, Bernstein, Palmore, & Hendin (1960) have conceptualized the psychotherapeutic interview as an informational-exchange system. Using this model, Siegman & Pope (1965, 1966) have isolated the effects of two important variables: (1) anxiety-producing topics and (2) ambiguity of interviewer's remarks.

These authors found that anxiety-arousing topics caused disrupted speech as measured by Mahl's (1956) speech-disturbance ratio (SDR) when "ahs" were not counted. This non-ah SDR was previously found to be an index of interviewee anxiety (Siegman & Pope, 1965). The second variable, ambiguous input, was found to produce cautious and hesitant speech from the interviewee rather than disruption (Siegman & Pope, 1965). The therapist's activity level was not found to be a relevant variable in the system (Pope & Siegman, 1962).

Despite the disparity listed above, both variables tended to result in an activation effect at mild levels of anxiety arousal. That is, anxiety-arousing topics were associated with greater vocabulary diversity as measured by the type-token ratio (TTR), fewer ahs, faster articulation rate, fewer silent pauses (Siegman & Pope, 1965), and greater productivity or number of words spoken (Feldstein & Jaffe, 1962). Ambiguous input activated only the lexical variables of vocabulary diversity (Siegman & Pope, 1966) and productivity (Pope & Siegman, 1965) but not necessarily such nonlexical or expressive components of speech as "ah," pauses, speech disturbances, etc.

From the above findings, it is obvious that both variables tended to activate the lexical components of productivity and vocabulary diversity during mild arousal, while only topical anxiety activated nonlexical components. One possible explanation for this difference might be that a person can more easily censor, modify, and process the lexical components (verbal diversity and productivity) of his speech as compared to the expressive or nonlexical components (ahs, speech rate, intonation, etc.). Moreover, the interviewee's perception of his own output (feedback) may produce differential effects on lexical and nonlexical production. For example, increased speech disruptions may increase the speaker's anxiety level, while increased vocabulary diversity may lower it. The instrumental role that the expressive factors play in anxiety production is still speculative (Mahl & Schultz, 1962).

The present experiment attempted to control nonlexical feedback of interviewees by having them write their responses. In this manner, productivity and vocabulary diversity could be investigated independently of the previously uncontrolled, feedback variable.

All of the investigations previously cited used only mild anxiety arousal. Duffy's (1962) curvilinear hypothesis of arousal postulates an activation effect at low levels of arousal but an inhibition effect at high levels. This study hypothesized that both productivity and vocabulary diversity would be inhibited under heightened anxiety according to this curvilinear position.

A second independent variable, response time, was additionally introduced. Since it has been a consistent finding that Ss carefully censor or measure their words in a two-person communication system (McGuire & Lorch, 1969), the amount of time available for such responding was investigated. It was hypothesized that vocabulary diversity would be significantly lower when censorship and word measuring was restricted by a limited response period.

Moderate arousal has been found to increase discriminatory powers (Kimble, 1961, p. 371). Maher (1966, p. 398) has pointed out that discrimination is reflected in language, "Thus, in a running sequence of writing or conversation, we might argue that the more limited the use of different words, the fewer are the discriminations that the patient is making." Extending the curvilinear hypothesis to verbal diversity, it would be expected that states of high arousal would be associated with low verbal discrimination and subsequently low TTRs. Without controlling interviewee feedback, Siegman & Pope (1965) found a significant order effect for topical anxiety, such that Ss rated their anxiety higher at the end of an interview when the anxiety-provoking topic was introduced first. That is, the topical anxiety resulted in *prolonged* anxiety reactions. Controlling nonlexical feedback, this investigation hypothesized that initial periods of high arousal would be associated with low TTRs due to the concomitantly lowered discriminatory powers of anxious Ss, but that the TTRs would systematically increase over time with the dissipation of anxiety. Such anxiety reduction over time was hypothesized in the absence of the sustaining effects of nonlexical feedback.

SUBJECTS

The Ss were 48 students enrolled in an introductory psychology course at Ohio University, with ages ranging from 17 to 25. All 23 males and 25 females were native speakers of English and were born in the United States. Although the experimental procedures did not match the Ss on intelligence, the random selection and assignment to experimental subgroups was considered sufficient control of this variable.

STIMULUS MATERIALS

Two cards from the Thematic Apperception Test were used as stimuli. Since the study wished to control the content of these stimuli with respect to the degree of affect suggested, two similar pictures applicable to both sexes were chosen. Card 7BM, which shows the heads and shoulders of two males, was shown to the male Ss, while 12F, similarly showing only the heads and shoulders of two females, was shown to the female Ss.

PROCEDURE

Four experimental groups were formed by the combinations of two levels of anxiety-arousing instructions and two different time periods in which to respond. The resulting groups were labeled: high anxiety-arousing instructions and 2 min (HAI-2), low anxiety-arousing instructions and 2 min (LAI-2), high anxiety-arousing instructions and 10 min (HAI-10), and low anxiety-arousing instructions and 10 min (LAI-10). Separate male and female randomizations were made so that the four subgroups were equally divided, with approximately six males and females in each. Each S was individually seated across from the E and given either anxiety-arousing or neutral instructions. The HAI were designed to arouse anxiety with respect to revealing information about the individual's intelligence, since ". . . This

picture was chosen because psychologists have discovered that the manner in which a person responds tells something about his intellectual ability. This is logical since good perception is associated with high intelligence." Those given LAI were told that the picture revealed whether or not they were affiliative in nature and were likely to join a fraternity or sorority. The previously hidden card was then handed to each S for his inspection. The E timed him with a concealed stopwatch from the time he received the card to the time he handed it back to the E. Each S was then instructed to write a story about the picture that told about (1) the identity of the people, (2) their motives, and (3) their emotions. Those in the HAI-2 and LAI-2 wrote their stories while being timed by the E, who had previously informed them that they would have only 2 min in which to respond. The HAI-10 and LAI-10 groups were sent into a second large room where they were timed by a second investigator.

PRODUCTIVITY AND TYPE-TOKEN RATIOS

The productivity is simply computed by counting the total number of words written. All contractions were counted as two words, and abbreviations were considered as the complete word. The TTR has the computational problem of controlling sample size, since the size of the ratio increases as a function of increasing sample size. The sample size can be controlled by dividing the speech corpus from each S into equal segments and by computing the mean segmental TTR (Johnson, 1944). Such a procedure has been shown to provide comparable TTRs for small samples of 25, 50, 75, and 100 words and has been previously used (Siegman & Pope, 1966). Those Ss who did not write a minimum of 25 words were excluded from the data when computing the TTR, but their data was retained for the rest of the analyses to avoid possible selection bias with respect to productivity.

RESULTS[1,2]

The mean number of words written for each of the four experimental conditions is illustrated in Fig. 1, where it is seen that the HAI served to reduce productivity as compared to the LAI. The 2-min condition limited productivity, as would be expected. These data were subjected to a 2 by 2 factorial design[3] that revealed significant differences in productivity for both the
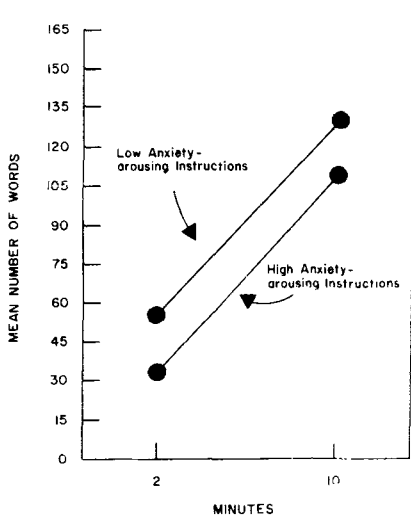

Fig. 1. Productivity as a function of high and low anxiety-provoking instructions and 2- or 10-min response intervals.

anxiety [F(1,44) = 7.19, p < .025] and the time [F(1,44) = 123.12, p < .001] factors. The interaction was nonsignificant.

The mean-segmental TTRs were similarly analyzed, with the resulting means reported in Table 1, which shows that lower TTRs were associated with the 2-min as compared with the 10-min condition. The 2 by 2 analysis of variance showed the time factor [F(1,42) = 8.76, p < .01] to be highly significant, while the anxiety factor was nonsignificant [F(1,42) = 1.22, p > .05].

In order to assess the hypothesized increase of TTRs over time, the first three segments of 25 words were analyzed in their consecutive order for the 10-min groups.[4] Thus, the first segment represented the earliest responses made by the Ss and the third obviously occurred later in time. By inspecting Table 2, it can be seen that the mean-segmental TTRs increased systematically over these three consecutive segments.

The data represented by the means in Table 2 were analyzed by a two-factor analysis of variance for repeated measures.[5] Here, the only significant factor was that of the segments [F(2,44) = 6.19, p < .01].

Table 1
Group Means for Mean-Segmental Type-Token Ratios

Time Period in Minutes	High-Anxiety Instructions	Low-Anxiety Instructions
2	.7440	.7800
10	.8247	.8419

Table 2
Mean-Segmental TTRs for 25-Word Segments

Groups		25-Word Segments 1	2	3
HAI-10	Mean-Seg. TTR	.80	.82	.84
	N	12	12	12
LAI-10	Mean-Seg. TTR	.74	.84	.84
	N	12	12	11
Combined Mean		.77	.83	.84
Combined N		24	24	23

DISCUSSION

The hypothesis that high levels of arousal are associated with an inhibition of productivity was clearly supported by the significantly fewer (p < .025) words written by Ss given HAIs as compared to LAIs. This finding gives indirect support to the curvilinear relationship hypothesized by Duffy (1962) between anxiety and behavioral variables. The highly significant difference between the productivity for the two different time groups is, of course, merely a function of the limited response period of the 2-min condition. The vocabulary diversity did not appear to be similarly inhibited as hypothesized. This output variable was significantly reduced, however, during the 2-min response periods, which apparently did not permit intensive censorship and word measuring. In summary, the data clearly indicated that anxiety arousal produced inhibited productivity, while reduced time in which to respond lowered verbal diversity; these two factors were statistically independent.

The systematic increase of TTRs over time supports the hypothesis that Ss' powers of discrimination increase when nonlexical feedback is not present to sustain anxiety reactions. This result could possibly be interpreted as a "practice effect." However, the significantly lowered TTRs for the 2-min groups suggests the former interpretation.

REFERENCES

DUFFY, E. Activation and behavior. New York: Wiley, 1962.

FELDSTEIN, S., & JAFFE, J. The vocabulary diversity of schizophrenics and normals. Journal of Speech & Hearing Research, 1962, 5, 76-78.

JOHNSON, W. Studies in language behavior: I. A program of research. Psychological Monographs, 1944, 56, 1-15.

KIMBLE, G. A. Hilgard and Marquis' Conditioning and learning. New York: Appleton-Century-Crofts, 1961.

LENNARD, H. L., BERNSTEIN, A., PALMORE, E. B., & HENDIN, H. The anatomy of psychotherapy. New York: Columbia University Press, 1960.

MAHER, B. Principles of psychotherapy. New York: McGraw-Hill, 1966.

MAHL, G. F. Disturbances and silences in the patient's speech in psychotherapy. Journal of Abnormal & Social Psychology, 1956, 15, 19-44.

MAHL, G. F., & SCHULZE, G. Psychological research in the extralinguistic area. In T. A. Seabeock, A. S. Hayes, and M. C. Bateson (Eds.), Approaches to semiotics. The Hague: Mouton, 1964.

McGUIRE, M. T., & LORCH, S. Natural language conversation modes. Journal of Nervous & Mental Disease, 1969, 146, 239-248.

POPE, B., & SIEGMAN, A. W. The effect of therapist activity level and specificity on patient productivity and speech disturbance in the initial interview. Journal of Consulting Psychology, 1962, 26, 489.

POPE, B., & SIEGMAN, A. W. Interviewer specificity and topical focus in relation to interviewee productivity. Journal of Verbal Learning & Verbal Behavior, 1965, 4, 188-192.

SIEGMAN, A. W., & POPE, B. Effects of question specificity and anxiety-producing messages on verbal fluency in the initial interview. Journal of Personality & Social Psychology, 1965, 4, 522-530.

SIEGMAN, A. W., & POPE, B. The effect of interviewer ambiguity-specificity and topical focus on interviewee vocabulary diversity. Language & Speech, 1966, 4, 242-249.

NOTES

2. When the time that the Ss looked at the stimulus card was compared, it was found that the HAI group (N = 24, $\bar{X}$ = 32.6 sec) looked significantly longer [F(1,46) = 7.73, p < .01] than the LAI group (N = 24, $\bar{X}$ = 23.7).

3. Two Ss were eliminated because they did not produce the minimal criterion of 25 words (HAI-2: N = 10).

4. A similar analysis was not possible for the 2-min groups since they produced too few words.

5. Since the combined mean for Segment 4 (N = 16, $\bar{X}$ = .86) and Segment 5 (N = 11, $\bar{X}$ = .87) had Ns that were significantly reduced, they were not included in the analysis so as to avoid giving disproportionate weight to their means.

A note on the "Search after meaning": Sensibleness of paraphrases of well formed and malformed expressions[1]

SAMUEL FILLENBAUM, University of North Carolina, Chapel Hill, N.C. 27514

Subjects were required to assess the sensibleness of sentences where each set of sentences included a source sentence plus some of its paraphrases. Characteristically, the paraphrases of malformed expressions (anomalous sentences, sentences with order inversions, word salads) were judged to be more sensible than their source expressions. This was taken as evidence for Ss' ability to impose sense on deviant expressions.

It seems reasonable to suppose that a listener will try to make sense of the materials he encounters, and that a malformed expression will, if at all possible, be "interpreted" in such a fashion as to make it intelligible. Linguists have offered various hypotheses to account for this ability to comprehend "semi-sentences" (see, e.g., Chomsky, 1961; Katz, 1964; Ziff, 1964). To the extent that a paraphrase task requires S to understand an utterance in order to construct another one equivalent in meaning, it provides a tool for externalizing that understanding, and may permit us to demonstrate this "search after meaning" in clear fashion. In the course of an exploratory study of paraphrasing, Ss were required to paraphrase both well formed expressions and expressions malformed in various ways. An examination of the paraphrases of the malformed sentences revealed that very often these were "improved," shifting to conventional form. Thus, mistakes in grammar were corrected, words in unconventional order were rearranged, incomplete phrases were completed, anomalous sentences were transformed into rather different but more sensible and intelligible versions, and various sorts of normalizing reshufflings improved word salads. The study to be reported below constitutes an attempt to objectify these impressions by having a new set of Ss rate the sensibleness of the paraphrases as well as that of their source sentences.

METHOD

We may distinguish two stages in the study that involved (1) obtaining of paraphrases, and (2) rating of the paraphrases and original expressions.

Obtaining the Paraphrases

A set of 48 sentences was constructed, including well formed expressions as well as semantically anomalous expressions, ungrammatical sentences, sentences with reversals in word order, incomplete sentences, and word salads. These sentences were arranged in random order and given to 40 Ss, college undergraduates, with the following instructions:

"Assume that you are trying to communicate each statement or phrase to someone else, and that it is essential that the other person get the sense of it as fully and exactly as possible. Your task is to paraphrase or rephrase each of the 48 items as accurately as you can, conserving every aspect of its meaning as completely as possible. In doing this, you may add or leave out words, you may rearrange or substitute words, as you find necessary."

This procedure yielded 40 paraphrases for each of the 48 source sentences. Next, 20 source sentences were selected in such fashion as to provide both well formed expressions and expressions malformed in various ways. There were seven ordinary, simple, well formed sentences, five semantically anomalous expressions, two expressions involving unusual order in word arrangement, three expressions constituting word salads, and three incomplete sentences. For each of these 20 source sentences, we now constructed two items, where each item consisted of the source sentence plus nine of its paraphrases, the paraphrases being selected and assigned at random to one or the other of the two items. The 10 sentences constituting each item were arranged in random order and typed on a single sheet; two sets of items were then assembled, the order of items being different in each set. Two sets were used, rather than one, to permit assessment of the replicability of the findings. The booklets representing the two item sets were arranged in alternation.

Rating of Source Sentences and Paraphrases

The Ss, 44 undergraduates of whom 21 responded to Set 1 and 23 to Set 2, were drawn from a population similar to that used in the paraphrase study. They were told:

"Most of the sentences in each set will be quite similar to each other. Your job is to assess the sensibleness or strangeness of the expressions as English (don't worry about the truth or falsity of what is said). For each set of items, read carefully through all 10 items a couple of times, and then pick out the three expressions which are most sensible, well formed, and acceptable as English, and put the letter A against each of these. Also pick out the three expressions which, as English, are the strangest or least acceptable, and put the letter Z against each of these. Remember to make your choice always relative to the 10 items in each set."

Given the way ratings were made, one can determine directly for each S, on each item, whether or not the source sentence was on the average judged more, as, or less acceptable than its paraphrases. If the source sentence is categorized as A, it is more acceptable than at least seven of the nine paraphrases, if it is categorized Z, it is less acceptable than at least seven of the nine paraphrases; otherwise, it is as acceptable as the paraphrases. Inspection of the frequencies of A and Z responses to the source sentence will thus indicate the extent to which it was judged more or less acceptable than its paraphrases, and a sign test on these frequencies will permit assessment of the significance of the differences obtained.