Abstract

The spontaneous speech of negative-syndrome schizophrenia patients is underproductive and contains many hesitations and pauses. Acoustic analysis of the patient's speech during interview reveals that the duration of pauses, independent of other linguistic or paralinguistic measures, correlates strongly with the clinician's impressions of the patient's flat affect and alogia. Pausing is less related to asociality and other aspects of the negative syndrome. The hesitations appear to reflect a word-finding difficulty that, together with neuropsychological evidence of compromised performance on word fluency tasks, suggests a specific speech generation difficulty. The significant relationship between pausing and both flat affect and alogia suggests that the two negative signs share phenomenal and psychometric properties. The examination of speech generation mechanisms may provide an informative avenue for study of schizophrenic psychopathology. Acoustic analysis reveals processes that are not apparent to the clinician and may provide a useful basis for clinical assessments and research.


Pauses

Not all pauses are dysfluent. Pauses may be purposeful, such as those used to dramatize the succeeding phrase or to provide meter or audible punctuation. A pause also may be the speaker's way of allowing the listener to catch up when the information load is great, reflecting the speaker's sensitivity to the listener's needs. However, the excessive pauses in the speech of negative-syndrome patients are not dramatic or emphatic, nor do they suggest sensitivity to the needs of the patients' conversational partners.

Pauses at different syntactic locations may reflect disturbances of specific speech generation processes. Switching pauses, which occur when the interviewer yields and the patient assumes the floor, reflect the time required to process the interviewer's question or comment and formulate a response. This reciprocal model presumes that the response depends on information in the interviewer's question or comment. However, some conversations, including some with thought-disordered patients, proceed more in parallel than as an interaction.

Pauses may also occur between clauses. At this location a pause can provide emphasis or assign importance to
Continuous speech from 5- to 10-minute semistructured interviews with normal controls, schizophrenia patients, and depressive patients. These results are calculated from an acoustic analysis (Alpert et al. 1989) of continuous speech from 5- to 10-minute semistructured interviews. The analysis includes all gaps that occurred during the patient’s turn. Analysis of continuous speech tends to find longer pauses than would spectrographic analysis of brief samples. The control group and the schizophrenia group were demographically similar; the depressive patients were a little older and included more females. Differences between the groups were noted for in-turn pauses \((F = 9.29; df = 2,53; p < 0.001)\) and for response latencies \((F = 8.23; df = 2,51; p < 0.001)\). Depressive patients showed increased pausing (see also Greden et al. 1981), but only in comparison with normal controls; the pauses for the schizophrenia patients were still longer. A number of other studies that used this system (Alpert et al. 1986) found similar group differences (Kring et al. 1994; Sison et al. 1996).

### Alogia and Flat Affect

Alogia has been related to both paucity of speech and impoverishment of speech content. The two characteristics are treated as alternate expressions of the same underlying process, although they differ in the amount of speech expected. We find that the training of raters for the assessment of impoverished speech content is problematic. The rater must take into account the patient’s education and IQ, allow for patient differences resulting from prolonged inpatient stays, and calculate the influence of other demographic factors that can produce artificial relationships between vacuous interests and psychopathology. Perhaps because of these problems, impoverished speech content is only weakly coherent with other measures of alogia.

Factor-analytic studies of ratings of combined positive and negative symptoms (Miller et al. 1993) and of only negative symptoms (Mueser et al. 1994), using different statistical approaches, found that paucity of speech and flat affect shared the same factor. This factor was different from that of impoverishment of speech content.

<table>
<thead>
<tr>
<th></th>
<th>In-turn pause (seconds)</th>
<th>Response latency (seconds)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal controls</td>
<td>20 0.9 0.21</td>
<td>1.5 0.50</td>
</tr>
<tr>
<td>Schizophrenia patients</td>
<td>19 1.6 0.77</td>
<td>3.4 1.93</td>
</tr>
<tr>
<td>Depressive patients</td>
<td>17 1.1 0.41</td>
<td>2.3 1.30</td>
</tr>
</tbody>
</table>

Note.—SD = standard deviation.
(SADS; Endicott and Spitzer 1978). We described these subjects previously (Alpert et al. 1993). Fifty-eight male and three female subjects were included from four sites, including three Veterans Affairs (VA) hospitals. Figure 1 presents the results of the regression of flat affect and alogia on in-turn pause time; figure 2 presents a regression of the two signs on the response latency measure. The regressions are highly significant and virtually parallel. The difference between the flat-affect and alogia scores is significant; the same average pause duration produces a higher flat-affect rating than alogia rating ($t = 4.34; n = 61; p < 0.001$). The two pause measures are correlated ($r = 0.63; n = 61; p < 0.001$), as are the ratings of flat affect and alogia ($r = 0.61; n = 61; p < 0.001$). In the same data set, the slopes and intercepts for ratings of asociality are quite different. While the pause measures are strongly associated with ratings of flat affect and alogia, other measures of speech productivity, such as the duration of utterances, the number of syllables per utterance or per total time, and speech rate during utterances, were less related to these clinical ratings, suggesting that it is the dysfluency rather than the speech underproductivity that provides the critical cue for the raters assessing the negative signs of flat affect and alogia.

Isolating the Effect of Pauses on Flat-Affect and Alogia Ratings

Converging evidence of the importance of pausing as a basis for clinical impressions of flat affect and alogia was found in a study in which pause durations were electronically altered (Alpert et al. 1995a). Other aspects of patients' speech were not changed. We assembled three rating tapes, each containing about 5 minutes from six recorded interviews. Three of the samples were left unaltered as filler samples; these were the same on each tape. Three other samples were edited to create alternate versions: an original, one in which the duration of each pause was doubled, and a third in which the duration of each pause was halved. Only one version of each sample was on a tape. The words, even the inflection and emphasis, were unchanged in all versions. Listeners could not detect that the recordings had been altered.

Clinicians each rated only one of the tapes, so each heard only one version of the three edited interviews. The version containing the lengthened pauses was rated as significantly flatter and more alogic (32% increase in flat-affect and alogia ratings); the short-pause version was rated less flat and alogic. This study provided specific evidence implicating pausing in clinical impressions of flat affect and alogia, although the results do not preclude the possibility that other cues also may contribute to clinical impressions.

For objective, behaviorally anchored rating scale items, acoustic measures of patients' speech can provide a yardstick against which to measure the validity of clinical assessment. For example, to evaluate assessments of the item “increased latency to respond,” which is included in both Andreasen's (1982) Scale for the Assessment of Negative Symptoms (SANS) and Alphs' Negative Symptom Assessment (NSA; Alphs et al. 1989), we mea-
Pausing and Thought Disorder

Much of the psychopathology research on pausing in schizophrenia has focused on the association between dysfluency and thought disorder. Thought-disordered schizophrenia patients tend to have longer between-clause pauses than do non-thought-disordered subjects (Rochester et al. 1977; Clemmer 1980; Resnick and Oltmanns 1984). These studies found that very long (greater than 5-seconds) between-clause pauses often are followed by incoherent speech; and it is specifically between-clause pauses that are associated with thought disorder. This pause type did not show a differential association with flat affect and alogia (Alpert et al. 1994).

Some disturbance in speech generation at the between-clause pause site (where the task is to formulate and carry forward the intended theme) might make the subsequent clause less cohesive and increase the risk of disorder. There is a circularity to this definition, and more direct and independent measures of impoverished speech and negative thought disorder are needed.

Neuropsychology of Fluency

It may be useful to contrast flat affect and alogia with nonfluent aphasia. While similar in some ways, flat affect and alogia and the aphasias differ in the quality as well as the quantity of speech behaviors. Nonfluent (Broca) aphasias are characterized by extremely sparse speech; considerable effort associated with the production of words; poor articulation; short phrase length (often a single word); loss of melodic, emphatic, and inflectional qualities of language (dysprosody); and a tendency to restrict speech to concrete content words (Benson and Geschwind 1985). Because of the monosyllabic output, pauses often are not found between words as in flat affect and alogia, where productivity is somewhat greater.

In flat affect and alogia, syntax is more complex and dysarthria is not characteristically present. Prosody, including affective prosody, which reflects the loss of movements associated with emotional expressiveness, and linguistic prosody, which involves word and phrasal stress, is reduced or absent (Alpert and Anderson 1977). This reduction in prosody probably interacts with dysfluency to limit productivity.

Discussing the neuropsychology of fluency, Lezak (1995) emphasized both ease and speed of verbal productivity, reflected in the length of uninterrupted strings of connected words. Neuropsychologists assess fluency by means of word-production tasks, counting the number of items (words) the subject produces (usually within a set time limit) within a restricted category (e.g., animals) or in response to a stimulus. Effortlessness is described as intrinsic to fluency but, to our knowledge, has not been operationalized. In an interesting aside to this issue, Sison et al. (1996) noted that flat-affect schizophrenia patients showed an increase in brow-corrugator electromyograph (EMG) activity during a task requiring them to describe emotional memories. Flat-affect patients generally show a decrease in activity of muscles involved in facial expression, but Darwin (1872/1965) had suggested that increased brow (forehead) activity accompanied mental effort. In the Sison study, the amount of pausing during an interview was a correlate of the amount of EMG corrugator activity during imagery. These results support the suggestion that ease of speech production is an important, albeit elusive, aspect of fluency.

Frith and Allen (1988) suggested that schizophrenia
patients have difficulty sustaining the lexical search process and hypothesized that different forms of this difficulty may underlie negative or positive thought disorder. Either the speech is sparse or it is repetitive and perhaps inappropriate. Both disturbances result from a defective self-directed search process. To test this, Allen et al. (1993) examined schizophrenia patients in a category-naming task. Their prediction, which the results confirmed, was that patients with poverty of speech would produce fewer items, while patients with incoherence would produce inappropriate (out-of-category) items. The authors also suggested that the reduced number of items did not simply reflect a smaller lexicon. It is possible to estimate the size of a subject's lexicon (i.e., the number of words in each category), by extrapolating from performance on a number of repetitions of the category-naming task. One can use the ratio of new items to the total number of items named on each iteration to calculate the size of the lexicon (Graesser and Mandler 1978). Lexicon size did not differ as a function of alogia. The size of a person's lexicon might be considered a correlate of IQ since the vocabulary subtest score is an indirect measure of lexicon size and is a strong correlate of full-scale IQ (Wechsler 1939). We could not find a significant relationship between the vocabulary subtest score and measures of pausing in flat-affect and non-flat-affect schizophrenia patients. Thus, the decreased speech productivity and hesitant speech of flat-affect and alogic patients appear to reflect some specific disturbance in speech generation.

Summary

We have presented evidence from several lines of research that suggests that there is a convergence between flat affect and alogia. The evidence also implies that both negative signs are related to a speech dysfluency. A disturbance in the mechanics of lexicon search, such as slower search speed, imprecision, or reduced resistance to distraction, might explain the low speech productivity and increased dysfluency of schizophrenia patients with negative signs. The syntactic location of many of the hesitations favors the hypothesis that specific rather than non-specific factors are at work.

It might be heuristic to explore the mechanisms relating extended response latencies and negative signs. Among the factors that might increase response latency are the following: (1) The patient requires extra time to decode the examiner's question or comment. For instance, a delay could result from a failure to recognize the interviewer's prosodic cues indicating end-of-turn intentions. (2) The cognitive operator calculating the response works more slowly in these patients. The switch (when one person yields the floor to another) is a busy transition, involving many operations; the switching pause could be extended where calculations are slowed. (3) The period during which responses are formulated may place cognitive operations at risk, and does appear to mark vulnerability of such operations during this period and would be consistent with the observed association between thought disorder and between-clause pauses. (4) There is a delay in the initiation of motor-response actions. Difficulty in initiation of motor acts is a hallmark of hypodopaminergic states and is consistent with the loss of emotion-associated movements that characterizes flat affect.

Other possibilities affecting response latency can be imagined on the basis of the unique processes that occur in the interactions between the interviewer and the patient. For such an explanation to be parsimonious, it should explain both the underproductivity of the patient's speech and its dysprosody. Although our data base of acoustic measures is insufficient to support specific hypotheses, focused data collection and cognitive studies with currently available methods could facilitate the design of experiments to discriminate among the possibilities.

This discussion was buttressed by results of acoustic analysis of speech. Determining the patterns and timings of pauses requires use of such acoustic methods; even highly skilled clinical listeners cannot discern these phenomena simply from listening to an interview. Acoustic methods provide a means of validating the clinicians' ratings, at least for objective, well-anchored items. When this validation is done, the intrusion of halo effects becomes clear. Raters can achieve reliability on the basis of interviews' shared halo variance while capturing very little of the specific item variance. Although reliability is necessary, it is not sufficient to establish validity. A clinical laboratory infrastructure, similar to that of clinical medicine, would be very useful for improving the precision of patient assessment in psychiatry. Acoustic measurement of the patient's voice is one possible tool for assessing disturbed feeling states and, perhaps, disordered thoughts.

References


The Authors

Murray Alpert, Ph.D., is Professor and Director of Psychology and Associate Director of the Millhauser Laboratories, Department of Psychiatry, New York University School of Medicine, New York, NY. Antonis Kotsaftis, Ph.D., is Clinical Instructor, and Enrique R. Pouget is Assistant Research Scientist, Millhauser Laboratories, Department of Psychiatry, New York University School of Medicine, New York, NY.

Erratum

The article by van den Bosch et al. entitled “What Determines Continuous Performance Task Performance?” (*Schizophrenia Bulletin*, 22(4):643–651, 1996) contains errors on pages 648 and 649. Figures 2 and 3 have been reversed; figure 2 that appears on page 648 should be figure 3 and figure 3 on page 649 should be figure 2. The captions and footnotes for the figures were placed in the article correctly.