Article

Vocal indicators of mood change in depression

ELLGRING, Heiner, SCHERER, Klaus R.

Abstract

In the framework of a major longitudinal study of depressive disorders, 11 female and 5 male depressives were audio-video-recorded while speaking with clinical interviewers. For selected utterances during depressed and recovered mood states several voice and speech parameters were obtained, using digital analysis techniques. As predicted, the results showed that an increase in speech rate and a decrease in pause duration are powerful indicators of mood improvement in the course of therapy (remission from depressive state). In female but not in male patients, a decrease in minimum fundamental frequency of the voice predicted mood improvement. These effects are discussed with respect to neurophysiological, cognitive, and emotional factors that have been suggested in the literature as possible causes for the patterns of motor expression observed in depressives. The data also point to the urgent need to systematically study gender differences in depressive speech behavior.

Reference

ELLGRING, Heiner, SCHERER, Klaus R. Vocal indicators of mood change in depression. *Journal of Nonverbal Behavior*, 1996, vol. 20, no. 2, p. 83-110

DOI : 10.1007/BF02253071

VOCAL INDICATORS OF MOOD CHANGE IN DEPRESSION

Heiner Ellgring and Klaus R. Scherer

ABSTRACT: In the framework of a major longitudinal study of depressive disorders, 11 female and 5 male depressives were audio-video-recorded while speaking with clinical interviewers. For selected utterances during depressed and recovered mood states several voice and speech parameters were obtained, using digital analysis techniques. As predicted, the results showed that an increase in speech rate and a decrease in pause duration are powerful indicators of mood improvement in the course of therapy (remission from depressive state). In female but not in male patients, a decrease in minimum fundamental frequency of the voice predicted mood improvement. These effects are discussed with respect to neurophysiological, cognitive, and emotional factors that have been suggested in the literature as possible causes for the patterns of motor expression observed in depressives. The data also point to the urgent need to systematically study gender differences in depressive speech behavior.

One of the major issues in the treatment of psychopathological disorder is the behaviorally based evaluation of progress in therapy. This problem is particularly troublesome in charting changes of mood during therapy of patients with affective illness. In general, mood evaluation scales and psychiatric diagnoses are used for this purpose. Unfortunately, neither of these approaches is entirely satisfactory, given the multiple sources of error, particularly in relation to patients' ability and willingness to communicate about their actual state and the difficulties in clinical inference of mood state on the basis of interview impressions. As a consequence, there has been renewed interest in the possibility of using objective behavioral

This study is based on psychiatric interviews that were conducted as part of the Longitudinal Study on Depression at the Max-Planck-Institute for Psychiatry, Munich. The authors acknowledge the contribution of Rudolf Bildhauer and Thomas Klos, who performed the vocal analyses (as part of their theses in partial fulfillment of the requirements for a degree in psychology) and who contributed to an earlier draft of this paper. Tom Johnstone and Marcia Smith provided valuable comments on the present manuscript.

Requests for reprints should be directed to: Prof. Klaus R. Scherer, Department of Psychology, University of Geneva, 9 route de Drize, CH-1226 Carouge-Geneva, Switzerland; Email scherer@uni2a.unige.ch.

measurement as an alternative for the assessment of mood changes across therapy (Maser, 1987).[1]

Given the well-established relationship between affect states and motor expression in body, face, and voice (Ekman, 1982; Scherer, 1979, 1986, 1989; Siegman, 1987a), it seems likely that the objective measurement of expressive behavior in several modalities might be useful in tracing changes in patients' mood over time. As shown in earlier reviews of this domain (Scherer, 1987), attempts to make use of acoustic measures of voice and speech measurement were used as early as the beginning of this century in order to evaluate the diagnostic value of vocalization for the study of depression (Isserlin, 1925; Scripture, 1921). Since then, there have periodically been renewed attempts to investigate this issue empirically (see Alpert, 1981; Darby, 1981; Jones & Pansa, 1979; Mahl & Schulze, 1964; Moses, 1954, Ostwald, 1964; Scherer, 1981, 1987; Siegman, 1987a). In this paper, we take a fresh look at recent work that has examined voice quality and temporal aspects of speech as indicators of changes in depressive patients' mood, and report a study that raises a number of issues with respect to the underlying mechanisms.

Temporal Aspects of Speech and Depression

It is generally expected that rate of speech will slow down in sad or depressed states and this is indeed what is found in virtually all of the studies on emotion encoding (portrayals by actors or laymen; see Pittam & Scherer, 1993, for a review). Furthermore, in studies in which depressed states are induced experimentally, rate or tempo of speech goes down (Markel, Bein, & Phillis, 1973; Natale, 1977; Teasdale & Fogarty, 1979). Given that depressed patients are expected to be depressed, one would expect them to speak at a very slow rate—which is indeed what clinical lore and psychiatric textbooks maintain.

In empirical research, reduced rate or tempo of spontaneous speech in depression is a frequently reported finding and seems to be a stable phenomenon (e.g., Hinchliffe, Lancashire, & Roberts, 1971; Mandal, Srivastava, & Singh, 1990; Newman & Mather, 1938; Pope, Blass, Siegman, & Raher, 1970; Weintraub & Aronson, 1967; see review by Siegman, 1987a). Because the temporal variables are not consistently defined, it is not clear to what extent speech rate is affected by pause time. The empirical studies that have focused on the number and duration of pauses in depressive speech have yielded somewhat inconsistent results, although there are a large number of studies that report pause lengthening in depressive state (Andreasen, Alpert, & Martz, 1981; Breznitz, 1992; Darby, Simmons, &

Berger, 1984; Kuny & Stassen, 1993; Nilsonne, 1987; Hinchliffe et al., 1971; Pope et al., 1970; Stassen, Bomben, & Günther, 1991; Vanger, Summerfield, Rosen, & Watson, 1992; Weintraub & Aronson, 1967). Some of the discrepancies in the findings in this area are due to widely varying criteria used in the definition of speech pauses (for example, time periods varying between two to five seconds have been used as criteria for the identification of silent pauses). Other factors, such as situational context, type of content, and type of speech sample, may also contribute to the discrepancies, and will be discussed in greater detail below.

On the whole, there is impressive evidence that depressive patients in acute and severe depressive states are likely to speak more slowly and with longer pauses. Therefore, an increase in speech tempo and shortening of pauses could be considered as promising candidates for markers of therapeutic success or remission from a depressive state. What is less clear is the mechanism that underlies this phenomenon. Three types of hypothesized mechanisms are generally encountered in the literature: persistent psychophysiological changes, cognitive impairment, and socio-emotional change.

The hypothesis assuming persistent psychophysiological changes in depressed patients generally assumes "psychomotor retardation" to be at the basis of the speech phenomena observed. In this tradition, pause time in non-spontaneous "automatic" speech (counting from 1 to 10) during depression has repeatedly been shown to be longer than during remission or in controls (Ghozlan & Widlöcher, 1988; Godfrey & Knight, 1984; Greden & Carroll, 1980; Greden, Albala, Smokler, Gardner, & Carroll, 1981; Hardy, Jouvent, & Widlöcher, 1984; Hoffmann, Gonze, & Mendlewicz, 1985; Szabadi, Bradshaw, & Besson, 1976; Teasdale, Fogarty, & Williams, 1980; see also Bouhuys & Mulder-Hajonides van der Meulen, 1984). The significant elongation of pause time while patients were depressed (compared to controls and to pause time after recovery) is interpreted as a marker for general psychomotor retardation or slowing. In five studies by Teasdale et al. (1980), using single case and group designs, a reliable positive correlation between self-rated depression and speech pause time was found for some but not for all of the patients studied. As patients in these studies did not exhibit clinical signs of retardation, Teasdale et al. (1980) hypothesized that pause time reflects variations in behavioral arousal. However, in a later study (Teasdale & Fennell, 1982) no effect on speech tempo was found.

In many of these studies a very high correlation (above $r = .9$) between speech rate and pause time was found. However, this could be due mainly to the specific task and to deducing speech rate from total duration for a standard task. Counting from 1 to 10 is different from spontaneous speech production, because it requires monosyllabic utterances without

involvement of complex cognitive planning processes. In this kind of "speech," temporal variation is likely to be almost exclusively due to pause time. In spontaneous speech, the number of pauses and the rate of articulation are additional sources of variance and the correlations between mean pause duration and speech rate are likely to be smaller (for example r = .74 in a study by Henze, 1954). Godfrey and Knight (1984) measured phonation time and pause time separately in the standard counting task. Although there was a significant effect of depression on pause time (which was found to correlate with total time of utterance at the level of r = .98), no effect on phonation time was found.

The use of automatic speech or counting in this area must therefore be regarded with some caution as it is hardly an ecologically valid sample of normal speech behavior. Even when used as a simple psychomotor task it is not ideal, because an acceleration in speech after recovery could occur due to practice effects from repeated measurements.

In any case, if psychomotor retardation is considered to be a general psychophysiological impairment affecting all motor behavior, possibly related to central noradrenergic deficiencies (Hoffmann et al., 1985) or a dysfunction of the mesolimbic-nigrostriatal system (Flint, Black, Campbell-Taylor, Gailey, & Levinton, 1993), the rate of articulation should be affected as much as if not more than pause time. Thus, in order to find conclusive evidence for the psychomotor retardation hypothesis one should find significant changes in both pause and phonation/articulation time, measured independently of total duration in natural speech tasks.

A second hypothesis that can be invoked to account for speech rate changes in depression postulates an impairment of cognitive functioning. Psycholinguistic work on temporal features of speech has shown that cognitive factors involved in speech planning and production are highly likely to affect rate and pausing (see Ballmer, 1980; Butcher, 1981; Boomer & Dittmann, 1962; Goldman-Eisler, 1961, 1968; Haenni, 1980; Rochester, 1973; Siegman, 1979, 1987; Siegman & Feldstein, 1979). It is not impossible, then, that cognitive changes in depressive states—with respect to attentional capacity or verbal content—are responsible for the temporal speech changes observed (see Siegman, 1987a, for an extensive discussion). Some authors postulate a combined effect of cognitive impairment and psychomotor slowing (Teasdale & Fennell, 1982). The changes in speech predicted under this hypothesis are mainly linked to the speech planning processes, i.e., hesitation pauses should increase in number and duration.

Finally, a social-emotional hypothesis can be formulated. Speech rate and pauses seem to be strongly affected by virtually all emotional states (Pittam & Scherer, 1993; Siegman, 1987b), including trait anxiety and state

anxiety (generally, speech pauses show a negative correlation to trait anxiety and a positive correlation to state anxiety, e.g., Kasl & Mahl, 1965; Siegman & Pope, 1972), as well as sad or depressed states (see above). As depression is characterized by habitual states of sadness and anxiety, one would expect to see the respective emotional speech rate changes in depression. In addition, Scherer (1979) has pointed out that variables like social anxiety and self-presentation need to be taken into account in studies on the function of speech pauses. The social interaction consequences of a depressed state have been described by a number of authors (Bouhuys & Alberts, 1984; Breznitz & Sherman, 1987; Hooper, Vaughan, Hinchcliffe, & Roberts, 1978; Vanger et al., 1992). The speech changes predicted from the vantage point of this hypothesis would be similar to what is observed under sadness (slowing of speech rate including more and longer pauses) or anxiety (slight increase in speech rate, not necessarily accompanied by a decrease in pauses).

A special variant of an emotion-based hypothesis has been proposed by Alpert and his collaborators, who suggested a strategy of identifying a particular syndrome, flat affect, and comparing the vocal and motor expression parameters for patients showing flat or non-flat affect, independent of nosological category. In other words, the flat affect syndrome, likely to be observable for parts of different patient goups (depressive, schizophrenic, etc.), is expected to cause the speech changes (rather than the specific pathology). A number of studies by this group of researchers have shown the utility of this distinction, finding shorter speech bursts, longer pauses, and less variability in fundamental frequency and energy of the voice for patients characterized by flat affect (Alpert, 1981; Alpert & Anderson, 1977; Andreasen et al., 1981). Levin, Hall, Knight, and Alpert (1985) have suggested that flat and inappropriate affect might be best defined in terms of poor differentiation between affects.

Fundamental Frequency of the Voice and Depression

Whereas the temporal variables of speech are determined by the sequence of motor commands, voice quality depends on the nature of the respiratory and phonatory characteristics (see Scherer, 1989). One of the most frequently studied voice variables is the fundamental frequency (F0) of the voice (subjectively perceived as voice pitch). This variable, which is directly related to the vibration of the vocal folds, is particularly interesting because of its relationship to the overall muscle tension of the speaker (see Scherer, 1979, 1986). Increased subglottal pressure and increased muscle tone tend to drive up F0. Paralinguistically, F0 variation is the major carrier

of prosodic information, responsible for the expressiveness of speech (Bolinger, 1978; Crystal, 1969; Fonagy & Magdics, 1963; Ladd, Scherer, & Silverman, 1985).

There are now quite a large number of studies reported in the literature that have investigated this parameter in relation to the clinical state of depression. Unfortunately, at least on the surface, the findings are not entirely in agreement. Most studies have reported a rather low mean F0 for depressives in relation to normals, or decreased F0 in an acute state of depression (Bannister, 1972; Eldred & Price, 1958; Moses, 1954; Roessler & Lester, 1976) although there are reports of an increase in F0 with the severity of depression (Whitman & Flicker, 1966). The opposite pattern was found in studies on vocal changes with therapy. Here, most studies have found a decrease in F0 after therapy or during positive mood states (Askenfelt & Sjoelin, 1980; Kuny & Stassen, 1993; Tolkmitt, Helfrich, Standke, & Scherer, 1982) whereas one study suggested an increase in F0 (Hargreaves & Starkweather, 1965). No significant effects were found by Nilsonne (1987) and Stassen et al. (1991).

The apparent discrepancy in the results may be explained at least in part by the fact that in most studies no clear distinction between manic and depressed states seems to have been made. Thus, it cannot be excluded that at least a certain percentage of the patients studied ought to have been diagnosed as bipolar. Circadian changes with improved states commonly occurring during the afternoon may play a role. Extreme care must be taken to analyze the manic and depressive phases of the patient separately, because opposite effects on F0 would be expected (see Scherer, 1987). Another possibility is that there are differential gender effects with respect to voice quality, particularly since pitch is an important secondary gender marker (Smith, 1979). This has not been systematically controlled in the literature so far.

Very consistent results have been found with respect to short-term changes in F0. In most studies, depressive speech has been found to have a rather narrow range as well as a restricted variability for F0 (Newman & Mather, 1938; Zuberbier, 1957; Ostwald, 1964; Hargreaves, Starkweather, & Blacker, 1965; Bannister, 1972). Accordingly, F0 range widens and variability increases after successful therapy and mood improvement (Kuny & Stassen, 1993; Nilsonne, 1987; Stassen et al., 1991). However, there are a few cases where the opposite effect, i.e., wide F0 range and variability, have been reported for depressive state (Fleischmann, 1980; Heitman, 1980).

Since the fundamental frequency of the voice is both an indicator of the physiological state of the individual and a paralinguistic signaling device (see Scherer, 1989), both the psychomotor retardation and the social-

emotional hypotheses outlined above can be invoked to explain the mechanisms underlying F0 changes in depression. It is difficult to invoke the cognitive hypothesis, since the effect of disturbances in cognitive performance on F0 level has been rarely examined. There is evidence, however, that increase in cognitive complexity of a task or other types of cognitive stress may increase F0 (see Griffin & Williams, 1987; Tolkmitt & Scherer, 1986).

With respect to physiology, a strong case can be made that changes in mood state during depressive illness are very likely to be accompanied by changes in fundamental frequency level and variability. The underlying mechanisms for this phenomenon are linked to the autonomic and somatic changes accompanying affective processes. Of particular importance is the strong link between fundamental frequency and muscle tension. On theoretical grounds (Scherer, 1979, 1987), supported by several pertinent findings in the literature, it can be predicted that depressive mood will lead to elevated muscle tone (Goldstein, 1965; Greden, Price, Genero, Feinberg, & Levine, 1984) with a consequent increase in F0 and a decrease of F0 range and variability (due to increasing rigidity of the phonation mechanism under high muscular tension). This rigidity would be consistent with the assumption of an impairment of the extrapyramidal system, as described by some researchers who support a psychomotor retardation hypothesis (e.g., Flint et al., 1993). During recovery F0 should decrease again and F0 range and variability should increase.

Aims of the Present Study

The review of the literature has shown that changes in both speech rate and F0 are promising candidates as markers of remission from depression after therapy. There is disagreement, however, as to the direction of the change to be expected (particularly for F0) and as to the underlying mechanism. The present study was conducted to help clarify the issues by attempting to replicate some of the earlier findings under more rigorous methodological conditions.

We will first briefly review the predictions of the three major types of hypotheses.

The psychomotor retardation hypothesis postulates a generalized impairment of the motor systems underlying speech behavior, independent of type of underlying emotion, gender of patient, type of speech sample used, or research context. Assuming extrapyramidally based effects, rigidity of the musculature should be the main agent, producing slower phonation

and articulation, imprecise of articulation, longer pause time, higher F0, and less F0 variation. These effects should disappear with appropriate medication.

The cognitive hypothesis invokes deficits in attentional and planning processes antecedent to speech behavior rather than impairment of motor behavior. Although these deficits might slow articulation, the major effect should be an increase in the number and duration of pauses, particularly hesitation pauses. There should be no effects on F0 caused by cognitive impairment alone.

The social-emotional hypothesis includes both cognitive and psychophysiological mechanisms, as the powerful effects of the dominant emotional state are expected to affect virtually all aspects of human functioning (see Scherer, 1986). However, contrary to the assumption of generalized psychomotor or cognitive impairment, the socio-emotional hypothesis would argue for a specificity of the observed effects with respect to the type of emotion underlying the depressed state. For example, if the underlying emotional state is resignation or apathy, one would expect slowed phonation/articulation, higher pause time, and lower fundamental frequency of the voice; if it is anxiety one would expect higher fundamental frequency of the voice with speech rate predictions varying depending on the type of anxiety (for example, speech rate should be slower if anxiety is accompanied by the feeling of helplessness; see Scherer, 1987).

The goal of the present study was to improve upon past research in three specific regards:

1) The explicit comparison of changes in the temporal aspects of speech and frequency aspects of voice. Only rarely have both types of parameters been assessed in the same study and even more rarely have they been systematically compared. However, the review of the rival hypotheses above show that differential predictions are made with respect to the overall patterning of speech and voice changes under depression and remission.

2) The explicit comparison of gender of patients. The large majority of patients studied in this area have been females. While there have been occasional references to gender differences (Andreasen et al., 1981; Darby, Hollien, & Lueck, 1977; Hooper et al., 1978; Long, 1988), the issue has not been addressed in a systematic fashion. This might be necessary to decide between hypotheses that postulate effects less likely to be affected by gender differences (such as impairment of the extrapyramidal system) versus those more likely to be affected by gender (such as the social-emotional hypothesis).

3) The explicit comparison of different topics in natural speech samples. As has been shown above, the nature of the speech activity requested

can strongly affect the speech behavior measured. This is very important because it may change the relationships between variables (e.g., pause time vs. articulation time), produce training effects, or prevent the evaluation of certain hypotheses (e.g., in the case where there is no cognitive planning required). The "automatic speech" method, asking patients to count from 1 to 10, is particularly problematic with respect to all of those problems. It is necessary, therefore, to use a type of speech sample that is ecologically valid and that allows all of the competing hypotheses to be properly evaluated. Since speech behavior is clearly affected by the topic discussed (Vanger et al., 1992), it is important to assess the relative effect of topic changes versus patient mood changes.

4) The explicit distinction between different types of temporal variables in speech. Much of the literature lacks a sufficiently precise operational definition of the variables. This is all the more problematic since the term "rate" can mean many things depending on how it is measured. Even more importantly, different types of pauses (e.g., articulation vs. hesitation pauses) may have very different functions and thus need to be clearly distinguished (see Scherer, 1982, on the problems of variable definition and operationalization).

5) The use of objective digital analysis techniques. In some work in the literature perceptual approaches to measurement were used; in others semi-objective (e.g., stop watch) assessment was used. Because the perception of temporal phenomena is highly context-dependent, it is preferable to automatize the measurement as much as possible.

6) The replication of findings to non-English speaking patients. Most of the research in this area has been conducted with patients from within a general Anglo-American culture, speaking various varieties of English. It is useful to examine the extent to which findings replicate across cultural and linguistic boundaries. This may be one way to test the psychomotor retardation versus the socio-emotional hypothesis since one would not expect any cultural differences if the phenomenon were to be due to biologically based psychomotor retardation.

Method

Subjects and Setting

Subjects were 16 hospitalized patients, eleven women and five men, with mean age of 55.6 years, participating in an interdisciplinary longitudinal study on depression (see Ellgring, Wagner, & Clarke, 1980; Ellgring & Ploog, 1985; Ellgring, 1989). Fourteen patients had been diagnosed as having endogenous depression (ICD 296.2, $n = 11$, 296.3, $n = 3$) by two

psychiatrists. Although patients were not diagnosed according to DSM-III-R (American Psychiatric Association, 1987), these diagnoses would correspond to Major Depression (296.2) and Bipolar Disorder, Depressed (296.5). Two patients, one male and one female, were classified as neurotic depression (ICD-9 300.4). According to DSM-III-R this would be Dysthymia or Depressive Neurosis. All patients were in a depressed state at the initial interview, and none was manic at remission.

Speech samples were taken from standardized interviews designed for studying the course of depressive illness (Standardized Interview for Depression, SID; see Ellgring, 1989). Interviews were conducted twice a week over the entire period of clinical treatment. From the total number of interviews, two were chosen for the analyses to be reported here: The first one during depression within Median = 5 days after admission to the hospital (henceforth called "depression"), and the second one during remission after Median = 50 days of clinical treatment (henceforth called "remission"). Care was taken to ensure that potentially biphasic patients were interviewed during a depressed phase for the first sample.

Subjective well-being was indicated by patients on a visual analogue scale (VAS), presented in the form of a mood thermometer (values were transformed to a range of 0 = "extreme well-being" to 100 = "extreme discomfort"). Across the total sample of 120 interviews in the study, the Visual Analogue Scale was correlated at $r = .74$ with a global clinical rating of depression and at $r = .84$ with an established scale of well-being ("Befindlichkeitsskala Bf-S"—State of Well-being Scale; von Zerssen, Köller, & Rey, 1970; see Ellgring, 1989, pp. 39-40, for further details on these correlations).

The interviews were carried out by three female and two male psychologists, as well as three male psychiatrists. Each patient was always interviewed by the same interviewer. In order to avoid effects of circadian mood changes, all interviews took place between 9:00 and 10:30 a.m. The first part of the interviews contained 28 standardized questions that covered psychosocial, emotional, and psychopathological aspects of depressive illness and lasted about 15 to 20 minutes. A "free dialogue" on specific problems of the patients of about 5 to 10 minutes followed.

Speech was recorded using Lavalier-type microphones for both patient and interviewer and a semi-professional Revox-A77 tape recorder.

Medication

Because of clinical needs as well as for ethical reasons, medication could not be balanced across states of depression and recovery. Medications administered were Amitryptiline (nine patients), Haloperidol plus Bi-

TABLE 1

Number of Female (F) and Male (M) Patients Receiving Different Types of Medication During Depressed and Recovered States

	Mood state	
Medication	Depressed	Recovered
Amitryptiline	2 F, 1 M	7 F, 1 M
Haloperidol plus Biperiden	1 F	1 F, 1 M
Lithium	1 M	1 M
Mianserin	1 F	1 F
No Medication	7 F, 3 M	2 F, 2 M

periden (two patients), Lithium (one patient), and Mianserin (one patient). Three patients received no medication throughout the clinical stay. Table 1 gives a detailed breakdown of the medication administered at the two points of measurement. The potential confounding effects of medication are raised in the Discussion section.

Selection of Speech Samples

The following sampling procedure was used for interviews during depressed and recovered mood states of each patient. The selection procedure allowed temporal sampling across the entire interview and across different topics. Four speech excerpts at two different points (T1, T2) during the standardized part of the interview were selected. Subsample T1 consisted of the answers to questions 1 and 2 ("Wie geht es Ihnen?"/"How are you?", "Wie war Ihre Stimmung überwiegend?"/How have you been feeling generally?"). Subsample T2 consisted of the answers to questions 27 and 28 ("Gibt es heute etwas, auf das Sie sich freuen?"/"Is there something you are looking forward to today?", "Wie glauben Sie, wird es auf längere Sicht mit Ihnen weitergehen?"/"How do you think you will be getting along in the future?"). Answers were required to be at least 25 syllables long for analysis. For 77% of all sampling points this criterion was met. In the 23% remaining cases the following or preceding answer was added. These answers were within 20 to 70 secs of the sampling point. In order to obtain a selection of utterances with sufficient stability according to speech rate, content, and time within the interview, the two answers for each sampling point were added together and treated as a single utterance.

As a third topic, a further subsample (T3) consisted of the first two an-

swers longer than 25 syllables in the free dialogue following the standardized interview. These answers were treated in the same way described above.

Voice and Speech Analysis

Speech. A verbatim transcript of the utterances was made. Each vocalization, including use of dialect, filled pauses, repeats, false starts, corrections, respiratory noises, and backchannel signals of the interviewer ("hmhm"), was noted. The following procedure was carried out using half speed of the tape-recorder (9.75 cm/sec instead of 19.5 cm/sec of the real time sound track). All backchannel signals which coincided with silent pauses of the patient were eliminated. There were no significant differences with regard to the occurrence of these instances when comparing the depressed versus the recovered state (chi square = 1.64, $p > .20$). If filled pauses in the patient's speech were preceded or followed by a longer silent period (criterion was 0.78 secs, i.e., 390 ms real time), the "filling signal" (ah, ehm, etc.) was deleted from the speech file subjected to digital pause analysis (see below). This procedure reduces a bias in digital pause detection as demonstrated elsewhere (Klos, 1984; Klos & Ellgring, 1987). As Boomer (1965) and Butcher (1981) argued, filled pauses are functionally similar to silent pauses. Butcher (1981, p. 90) gives empirical evidence that durations of filled pauses do not differ significantly in their distribution from silent pauses.

Using digital pause detection and the correction procedures described above, the following speech parameters were extracted: speech rate (SR), mean pause duration (MPD), and number of pauses (NOP). SR was calculated by dividing the number of syllables in the transcript by length of utterance in minutes (i.e., syllables per minute). MPD was obtained by averaging all pauses longer than 390 ms within utterances. NOP is the number of pauses longer than 390 ms per 100 syllables. A detailed description of the analyses can be found in Klos (1984).

Voice. Analog speech signals were digitized using a sampling frequency of 16384 Hz and a filter frequency of 6553 Hz. Fundamental frequency (F0) was extracted after elimination of all silent periods longer than 240 ms. Software of the Giessen Speech Analysis System (GISYS) was used (see Scherer, 1982; Standke, 1992; Tolkmitt et al., 1982). One patient (female) had to be excluded from this analysis because of excessive background noise on the tape. The following voice parameters were extracted: mean F0 (MF0), range of F0 (RaF0), and minimum F0 (MinF0). A detailed description of the analyses can be found in Bildhauer (1985).

TABLE 2

Intercorrelations Between Voice and Speech Parameters

Parameter	RaF0	MinF0	SR	NOP	MPD
MF0	.51***	.84***	−.14	.14	.18
RaF0		.04	.02	−.06	−.02
MinF0			−.27	.27	.27
SR				−.73***	−.54***
NOP					.13

Note. MF0 = Mean F0, RaF0 = Range F0, MinF0 = Minimum F0, SR = Speech Rate, NOP = Number of Pauses, MPD = Mean Pause Duration. *$p < .05$, **$p < .01$, ***$p < .001$

Results

Mood Ratings

There were major changes in subjective well-being following therapy (see Figure 1). Values on the Visual Analogue Scale improved for each of the patients (sign test, $p < .01$, two-tailed). Median values improved from Median = 75 during depression to Median = 21 at remission. As can be seen in the figure, the extent of the improvement was generally higher for the female patients.

Differences Between Topic-Related Speech Subsamples

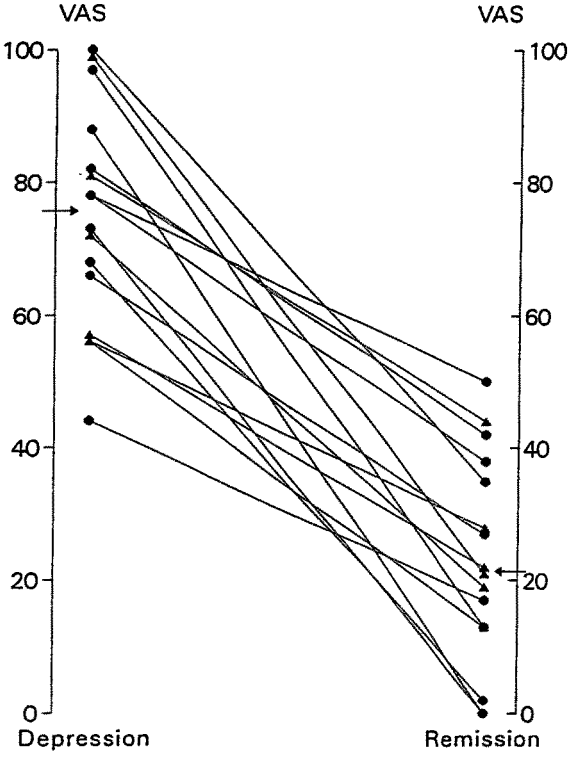
Mean length of the three speech subsamples (T1, T2, T3) was 100.6, 95.9, and 103.7 syllables during the depressed state and 109.4, 109.7, and 126.9 syllables during the recovered state. The differences between sampling points were not statistically significant, and there was no significant interaction between sampling point and topic (Wilcoxon test, two-tailed, $p > .10$). Since the intercorrelations of the voice and speech variables between the different sampling points of the speech samples (T1, T2, T3) were very high and did not differ significantly within each variable (r = .86, .88, and .94 for mean F0; r = .94, .95, and .94 for minimum F0; and r = .79, .78, and .71 for speech rate), T1, T2, and T3 were averaged to obtain a single value per parameter for each patient in both states. The lack of significant differences between the three subsamples indicates that neither the differences in topic (content) nor between form (standardized inter-

view vs. free dialogue) affected the speech and voice parameters under measurement.

Intercorrelations of Voice and Speech Parameters

Pearson *rs* were computed over all available sampling points (32 values, depressed and recovered mood states for 16 patients). These correlations are shown in Table 2. For voice parameters, Mean F0 was closely correlated with Minimum F0 ($r = .84$) and Range of F0 ($r = .51$). Range and Minimum of F0, however, appeared to vary independently ($r = .04$). For speech parameters, Speech Rate showed a strong negative relationship with Number of Pauses ($r = -.73$) and Mean Pause Duration ($r = -.54$),

Figure 1. Subjective well-being of depressed patients. Individual values obtained via self ratings on a Visual Analogue Scale (VAS) during depression and remission. High values: negative well-being, low values: positive well-being. ○ = female, ▽ = male patients. Median values indicated by arrows (→).

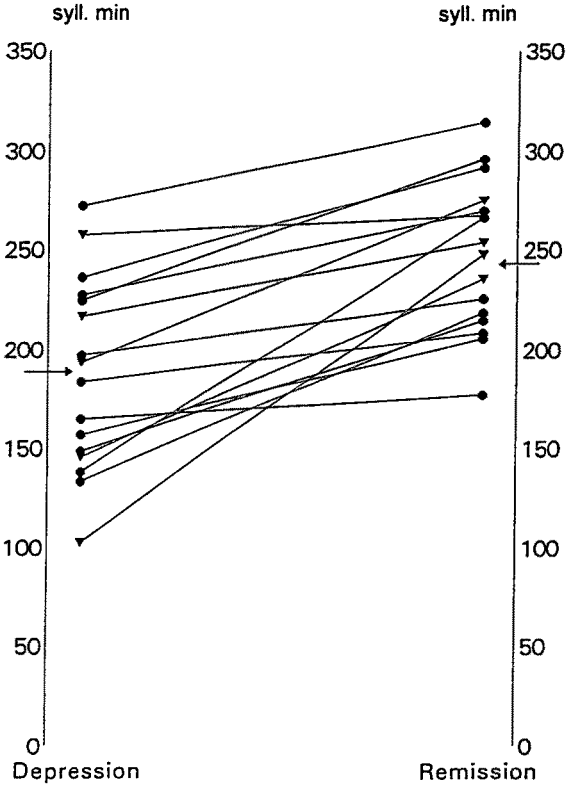


whereas no substantial relationship existed between Number of Pauses and Mean Pause Duration (r = .13). Voice and speech parameters appeared to vary independently as shown by the absence of significant correlations between the two types of parameters.

Therapy Effects on Speech

Speech rate (SR) increased noticeably from depressed to recovered states in all patients. This effect could be observed for both male and female patients (see Figure 2). Female patients accelerated about 22.6% from 189.1 to 244.4 syllables per minute and male patients accelerated 24.0% from 194.2 to 255.7 syllables per minute (Wilcoxon test, two-tailed, p = .003 and .04, respectively).

Figure 2. Speech rate of depressed patients. Individual values (syllables/minute) for female (○) and male (▽) patients during depression and remission. Medians = →.

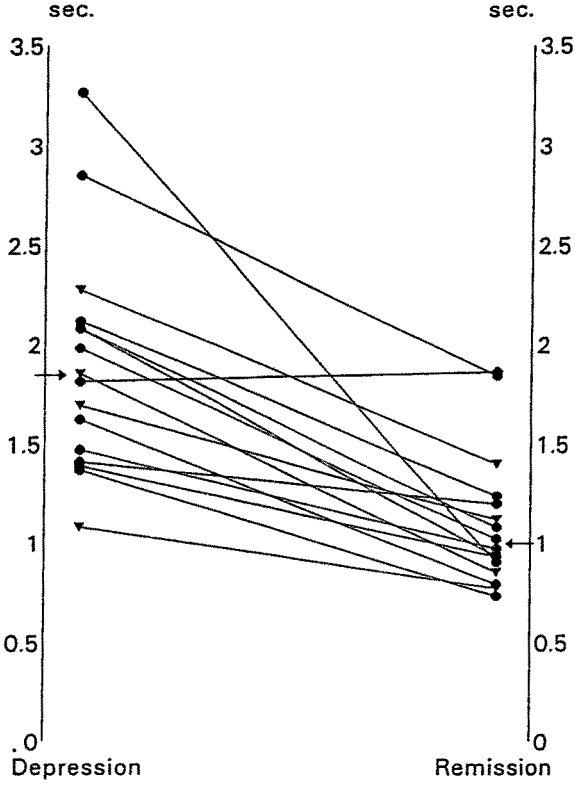


Mean pause duration decreased about 59% from an average of MPD = 1.94 secs to MPD = 1.14 secs for female and 56% from MPD = 1.80 secs to MPD = 1.01 secs for male patients (Wilcoxon test, two-tailed, p = .004 and .04; see Figure 3).

The average number of pauses decreased about 25% in female patients (NOP = 7.89 to NOP = 5.88 pauses per 100 syllables, p = .05) and about 32% in male patients (from NOP = 7.58 to NOP = 5.12/100 syllables, $p < .07$).

Clearly, the results reported above are not independent of each other. Since speech rate can be expected to be determined in large part by paus-

Figure 3. Mean pause duration (MPD) of depressed patients. Individual values (secs) for female (o) and male (∇) patients from speech samples during depression and remission. Medians = →.



ing (as confirmed by the correlations in Table 2), the results may well be redundant. This was confirmed by regressing MPD and NOP on SR and analyzing the residuals for potential effects of mood improvement. Paired *t*-tests of the residuals did not yield any significant effects, suggesting that the SR results are due almost exclusively to pausing rather than to changes in articulation/phonation rate.

Therapy Effects on Voice

As shown above, F0 mean and minimum F0 correlated highly. It was thus decided to use only one of these indicators. As shown by Scherer, Ladd, and Silverman (1984), minimum F0 can be considered a more adequate measure of the physiological baseline level of F0, independent of speech-induced variations and of outlier values. In the present case, this is corroborated by the fact that MinF0 is more independent of F0 range than MF0. Therefore, MinF0 was chosen for the analyses of mood changes.

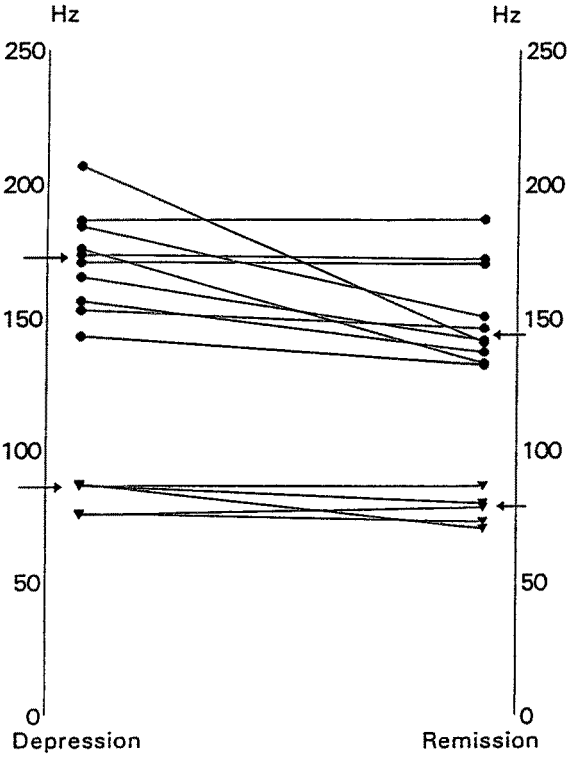
The results for MinF0 are shown in the form of individual data plots indicating change from depressed mood to remission in Figure 4. It should be noted that F0 shows a characteristic difference of more than 100 Hz between male and female speakers (Smith, 1979). Therefore, F0 measures are not directly comparable between male and female patients.

As one would expect after inspection of the individual data points, no significant differences in the F0 variables between depressed and recovered moods were found for men. However, significant changes were found for F0 minimum and F0 range between depressed and recovered phases in women. For this group, there was a decrease of 20 Hz for average F0 minimum (*t*-test, $p = .016$, two-tailed) and approximately the same amount of increase for F0 range ($p = .03$, two-tailed) from depressed to recovered mood.

Correlations Between Speech/Voice Variables and Subjective Well-Being

Given the differences in mood improvement and in voice changes for men and women, it was decided to compute the correlations separately for men and women. As shown in Table 3, for women, MinF0, SR, and MPD were significantly correlated with subjective well-being. The correlations with RaF0 and NOP were marginally significant ($p < .06$). For the male patients, on the other hand, no significant correlations were found (although there is a similar tendency in the data for the speech but not the

Figure 4. Minimum fundamental frequency (MinF0) of depressed patients. Individual values (Hz) for female patients (○) and male patients (▽) from speech samples during depression and remission. Medians = →.



voice variables). These gender differences could be due to the fact that the male patients did not improve as much over therapy, thereby reducing the variance in the data, and/or different mechanisms underlying the relationship between mood and speech/voice.

Discussion

We first discuss two potential confounding factors: order and medication effects.

TABLE 3

Correlations Between Voice/Speech Parameters and Subjective Well-being

	Subjective well-being	
	VAS female *df*=22	VAS male *df*=10
RaF0	−.43	.01
MinF0	.69***	.28
SR	−.61**	−.58
NOP	.41	.59
MPD	.72***	.49

Note. MF0 = Mean F0, RaF0 = Range F0, MinF0 = Minimum F0, SR = Speech Rate, NOP = Number of Pauses, MPD = Mean Pause Duration, VAS = Visual Analogue Scale for Subjective Well-being. *$p<.05$, **$p<.01$, **$p<.001$

Order Effects

Serial order of interviews and state of the patients are confounded. However, there are two arguments against the assumption of effects of serial order on results concerning temporal speech patterns. First, a comparison of first and second interviews (n = 9) to later interviews in depressed states (third to eighth interview, n = 7) yielded no significant differences in SR during the standardized part of the interview (T1, T2) (Early M = 186.7 syllables/sec, SD = 35.5 vs. Late M = 251.5 syllables/sec, SD = 53.6; t (14) = .78, $p > .20$, two-tailed). Second, T3 was a free dialogue, varying with each interview. As mentioned above, T3 did not differ from T1 and T2. In consequence, serial order effects are unlikely to have had a strong effect on the voice and speech changes found.

Medication Effects

For the following reasons, the effects of medication are likely to be small or point to a direction of change opposite to our hypotheses.

The neurotically depressed patients received no medication in this study. The remainder received either neuroleptic or tricyclic substances at

different points during the course of therapy. Neuroleptic substances affect the extrapyramidal motor system and thus may have a slowing effect on speech. However, in endogenously depressed patients the therapeutic effect of medication is not expected to occur earlier than 10 days after administration. There might be earlier effects on speech production but there is no evidence of this in the literature. For two patients who were given neuroleptic substances, we balanced medication across both interviews. Another patient was receiving Haloperidol at the time of the recovered state interview, which is when we expected speech to be accelerated.

Most of the subjects had Amitryptilin during the recovered state only. Those tricyclic antidepressants have been shown to slow down temporal patterns of speech (e.g., Schatzberg, Cole, & Blumer, 1978). Our hypothesis, however, predicts accelerated speech, that is normal speech, during the recovered mood period.

Thus, "medication effects," i.e., a slowing down of speech, would be expected to work against "recovery effects" in the recovered state and attenuate possible differences. Therefore, medication effects should work against our hypothesis. Effects of the medication on voice F0 are difficult to assess. The literature provides no answer to this question. Post-hoc comparisons of changes in voice parameters did not point to systematic changes due to medication.

Replication of Earlier Findings

The present results replicate and extend the findings in the literature despite the fact that they were obtained with German-speaking patients in a German hospital and with a somewhat different methodological approach with respect to speech sampling and analysis. As far as temporal speech parameters are concerned, the speech of both male and female depressives seems to be slowed in the acute state as inferred from the significant increase in rate with mood improvement at remission. While articulation/phonation times were not separately measured in this study, the high correlations of speech rate (and of subjective well-being) with the pause parameters seem to indicate that much of the effect is due to a reduction in the number and the duration of pauses. This is confirmed by the very weak and insignificant differences between SR residuals (with NOP and MPD effects removed) before and after therapy.

As far as vocal frequency is concerned, the present data confirm earlier studies that found, mainly for female patients, higher F0 in acute depression (and F0 decrease with remission) rather than the opposite pattern of findings (which might have been partly due to the lack of separation

between depressed and manic states in bipolar patients). This pattern is not found for male patients. However, since male depressives have been rarely studied in the literature, it is difficult to compare the present findings—which are again based on a very small sample—with earlier data.

Underlying Mechanisms

The systematic comparison of temporal speech and voice frequency parameters as well as suggestive gender differences found in this study may help to narrow the options and to suggest possible designs for future studies in this area.

Although the results of the present study are not incompatible with a psychomotor retardation hypothesis, there are several aspects that mitigate against the assumption that this hypothesis, by itself, can explain the data. To begin with, if speech motor behavior in depression is affected by neurotransmitter disorders or other types of neurological dysfunctioning, one would not expect to see any gender differences. More importantly, if efferent motor commands are consistently impaired one would expect a very general effect of muscular rigidity on speech behavior, including slowing and greater imprecision of articulation. However, in this study and in some of the earlier studies that attempted to systematically evaluate the determinants of changes in speech rate, pausing seems to have a much greater effect than articulatory slowing. Articulatory precision was not studied in this research. There is some evidence for articulatory imprecision in acutely depressed states (Flint et al., 1993; Tolkmitt et al., 1982). However, these effects can also be explained by increased tension of the articulatory musculature due to emotional effects (see Tolkmitt et al., 1982; Tolkmitt & Scherer, 1986). Therefore, while psychomotor retardation, possibly caused by neurological disorders, may explain or contribute to the effects in some individuals, by itself it can hardly account for the effects found in the literature.

Similarly, an exclusively cognitive hypothesis seems rather unsatisfactory as an account of the underlying mechanisms. Again, it is rather difficult, on the basis of that hypothesis, to account for the suspected gender differences. Furthermore, one might have expected at least some differences due to the type of speech situation (standardized interview with simple questions vs. free discussion) or for type of topic. Most importantly, an exclusively cognitive explanation does not provide much justification for the massive changes in F0 that were observed for the female patients.

The hypothesis that seems best able to accommodate the findings in this study as well as in much of the literature is the socio-emotional explanation. As mentioned above, emotional states, including affect distur-

bances, are expected to be accompanied by cognitive and physiological changes, thus covering some of the domains of the competing hypotheses. However, rather than assuming general impairments, the social-emotional approach assumes the effects of depression on speech and voice to correspond to the major underlying type of emotion. Clinical literature as well as empirical research on the facial expression of emotion (Ellgring, 1986) suggest that sadness and dejection, anxiety, or anger are all possible emotions underlying depression. Depending on the dominant emotion, different patterns of effects on speech and voice are predicted (Scherer, 1986, 1987). Consequently, individual reactions or group differences (such as gender effects) can be conceptualized and studied as effects of different underlying emotions.

The present data, while inconclusive due to the very small number of male patients, suggest an intriguing possibility that might be worth further exploration in future research. The results for the female patients can be rather well explained by the assumption of underlying anxiety states that are reduced through therapy. As mentioned in the introduction, an increase in muscle tension (driving up F0) is one of the effects expected for states of stress, anxiety, or fear. Furthermore, relationships between anxiety and speech pauses have been generally postulated for normal subjects (see Murray, 1971), and one might expect similar findings for depression, given that anxiety is a symptom frequently associated with depression, particularly for female patients (Long, 1988; Teasdale et al., 1980). Since state anxiety seems to increase pausing (see Introduction), this would explain the general slowing of speech in the female patients.

The results for male patients cannot be easily explained by anxiety due to the absence of clear voice frequency effects. However, the slowing of speech under depression (and the speeding up after remission) could be due to resignation, sadness, or dejection—possibly due to the feeling of hopelessness or powerlessness—in a state of acute depression.

Future Research

Concretely, an intriguing hypothesis suggested by the present data is that female depressive patients may often have anxiety as the central emotion underlying their depressive state (possibly due to worrying about being unable to cope or helpless in specific life situations) whereas male patients might be characterized by underlying sadness and resignation (possibly due to repeated feelings of powerlessness in situations where they are expected to perform well). Given the different etiologies of the depressive state for different genders, the effects on voice and speech are likely to be somewhat different.

Obviously, a systematic investigation of this hypothesis would require a research design and methodology that surpasses the habitual standards in the field. Apart from a sufficiently large sample of male and female depressive patients with homogeneous nosological diagnosis, controlled therapeutic intervention, systematic recording of ecologically valid speech samples of various kinds (including variations of topic, dialogue form, and cognitive demands), as well as state-of-the-art speech and voice analysis (including precise operationalization of the parameters to be measured), a large number of factors would need to be assessed to allow to evaluate the different hypotheses: assessment of emotional state (using self-report and facial expression analysis), effects of medication, evaluation of neurotransmitter function and neurological status, individual differences (personality, coping strategies), position in social networks, as well as interaction and communication strategies (given the strategic use of voice and speech parameters). Given the reality of clinical research settings, these are clearly unrealistic demands. However, major progress in this area is unlikely to come about unless at least some of the desiderata outlined above can be fulfilled.

Clinical Application

These findings confirm the usefulness of objective expression measurement in providing important information for diagnosis and evaluation of therapy. Vocal variables can be easily obtained using normal interviewing procedures, and a large number of speech and voice variables can be assessed with standardized and reliable procedures. In addition, digital speech analysis techniques, allowing, for example, the extraction of fundamental frequency of the voice from a tape recording of a patient's utterance, are becoming increasingly available on small microcomputers (see Scherer, 1982, 1989, for a review of methodological procedures). Therefore, it would seem that the attempt to trace changes in mood state by using objective behavioral indicators is likely to yield important information for the assessment of affective illness both in terms of research and clinical practice.

Note

1. Renewed in the sense that early psychiatrists had considered expressive manifestations of affective illness as a central aspect of clinical description and diagnosis (e.g., Bleuler, 1950; de Sanctis, 1904; Kraepelin, 1921). This interest in behavioral symptoms was somewhat lost during the heyday of so-called "objective clinical assessment."

References

Alpert, M., & Anderson, L. T. (1977). Imagery mediation of vocal emphasis in flat affect. *Archives of General Psychiatry, 34,* 208-212.

Alpert, M. (1981). Speech and disturbances in affect. In J. K. Darby (Ed.), *Speech evaluation in psychiatry* (pp. 359-368). New York: Grune & Stratton.

American Psychiatric Association (1987). *Diagnostic and statistical manual of mental disorders* (3rd ed. revised). Washington, DC: American Psychiatric Association.

Andreasen, N. C., Alpert, M., & Martz, M. J. (1981). Acoustic analysis. *Archives of General Psychiatry, 38,* 281-285.

Askenfelt, A., & Sjoelin, A. (1980). Voice analysis in depressed patients: Rate of change in fundamental frequency related to mental state. *Speech Transmission Laboratory—Quarterly Progress and Status Report. Stockholm, 2-3,* 71-84.

Ballmer, T. T. (1980). The role of pauses and suprasegmentals in a grammar. In H. W. Dechert & M. Raupach (Eds.), *Temporal variables in speech* (pp. 211-221). The Hague: Mouton.

Bannister, M. L. (1972). *An instrumental and judgemental analysis of voice samples from psychiatrically hospitalized and nonhospitalized adolescents.* Unpublished doctoral dissertation, University of Kansas.

Bildhauer, R. (1985). *Vokale Indikatoren depressiver Stimmung.* (Vocal indicators of depressed mood). Unpublished diploma thesis, University of Giessen.

Bleuler, E. (1950). *Dementia praecox or the group of schizophrenias.* (Orig. Ed. 1911). New York: International Universities Press.

Boomer, D. S. (1965). Hesitation and grammatical encoding. *Language and Speech, 8,* 148-158.

Boomer, D. S., & Dittmann, A. T. (1962). Hesitation phenomena in spontaneous English speech. *Word, 15,* 19-44.

Bolinger, D. L. (1978). Intonation across languages. In J. P. Greenberg (Ed.), *Universals of human language . Vol. 2. Phonology.* (pp. 471-524) Stanford: Stanford University Press.

Bouhuys, A. L., & Alberts, E. (1984). An analysis of the organization of looking and speech-pause behaviour of depressive patients. *Behaviour, 89,* 269-298.

Bouhuys, A. L., & Mulder-Hajonides van der Meulen, W. R. E. (1984). Speech timing measures of severity, psychomotor retardation, and agitation in endogenously depressed patients. *Journal of Communication Disorders, 17,* 277-288.

Breznitz, Z., & Sherman, T. (1987). Speech patterning of natural discourse of well and depressed mothers and their young children. *Child Development, 58,* 395-400.

Breznitz, Z. (1992). Verbal indicators of depression. *The Journal of General Psychology,* 119, 351-363.

Butcher, A. (1981). Aspects of the speech pause: Phonetic correlates and communicative functions. *Universität Kiel, Arbeitsbericht Nr. 15 des Instituts für Phonetik.* (Reports of the Institute of Phonetics of the University of Kiel).

Crystal, D. (1969). *Prosodic systems and intonation in English.* Cambridge, UK: Cambridge University Press.

Darby, J. K., Hollien, H., & Lueck, P. (1977). Vocal and speech patterns of depressive patients. *Folia Phoniatrica, 29,* 279-291.

Darby, J. K. (1981). Speech and voice studies in psychiatric populations. In J. K. Darby (Ed.) *Speech evaluation in psychiatry* (pp. 253-284). New York: Grune & Stratton.

Darby, J. K., Simmons, N., & Berger, P. A. (1984). Speech and voice parameters in depression: A pilot study. *Journal of Communication Disorders, 17,* 75-85.

Ekman, P. (Ed.). (1982). *Emotion in the human face* (2nd ed.). Cambridge, UK: Cambridge University Press.

Eldred, S. H., & Price, D. B. (1958). The linguistic evaluation of feeling states in psychotherapy. *Psychiatry, 21,* 115-121.

Ellgring, H., Wagner, H. & Clarke, A. H. (1980). Psychopathological states and their effects on speech and gaze behaviour. In H. Giles, W. P. Robinson, & P. M. Smith (Eds.), *Language: Social psychological perspectives* (pp. 267-273). Oxford: Pergamon Press.

Ellgring, H., & Ploog, D. (1985). Sozialkommunikatives Verhalten in klinischer Perspektive. (Socio-communicative behavior in clinical perspective). In D. Bente, H. Coper & S. Kanowski (Hrsg.), *Hirnorganische Psychosyndrome im Alter II* (pp. 217-236). (Organic brain syndromes in aging). Berlin Heidelberg: Springer.

Ellgring, H. (1986). Nonverbal expression of psychological states in psychiatric patients. *European Archives of Psychiatry and Neurological Sciences, 236*, 31-34.

Ellgring, H. (1989). *Nonverbal communication in depression.* Cambridge, UK: Cambridge University Press.

Fleischmann, G. (1980). *Akustisch-phonetische Untersuchungen an der Sprache der Depression.* (Acoustic-phonetic studies of the speech of depressives). Unpublished master's thesis, University of Cologne.

Flint, A. J., Black, S. E., Campbell-Taylor, I., Gailey, G. F., & Levinton, C. (1993). Abnormal speech articulation, psychomotor retardation, and subcortical dysfunction in major depression. *Journal of Psychiatric Research, 27*, 309-319.

Fonagy, I., & Magdics, K. (1963). Emotional patterns in intonation and music. *Zeitschrift für Phonetik, 16*, 293-326.

Ghozlan, A., & Widlöcher, D. (1988). Ralentissement psychomoteur depressif et traitement de l'information. (Depressed psychomotor retardation and information processing). *Psychologie Francaise, 33*, 251-258.

Godfrey, H. P. D., & Knight, R. G. (1984). The validity of actometer and speech activity measures in the assessment of depressed patients. *British Journal of Psychiatry, 145*, 159-163.

Goldman-Eisler, F. (1961). The significance of changes in the rate of articulation. *Language and Speech, 4*, 171-174.

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech.* London: Academic Press.

Goldstein, I. B. (1965). The relationship of muscle tension and autonomic activity to psychiatric disorders. *Psychosomatic Medicine, 27*, 39-52.

Greden, J. F., Albala, A. A., Smokler, I. A., Gardner, R. R., & Carroll, B. J. (1981). Speech pause time: A marker of psychomotor retardation in endogenous depression. *Biological Psychiatry, 16*, 851-859.

Greden, J. F., & Carroll, B. J. (1980). Decrease in speech pause time with treatment of endogenous depression. *Biological Psychiatry, 15*, 575-587.

Greden, J. F., Price, H. L., Genero, N., Feinberg, M., & Levine, S. (1984). Facial EMG activity levels predict treatment outcome in depression. *Psychiatry Research, 13*, 345-352.

Griffin, G. R., & Williams, C. E. (1987). The effects of different levels of task complexity on three vocal measures. *Aviation, Space, and Environmental Medicine, 58*, 1165-1171.

Haenni, R. (1980). What is planned during speech pauses? In H. Giles, W. P. Robinson, & P. M. Smith (Eds.), *Language: Social psychological perspectives* (pp. 321-326). Oxford: Pergamon Press.

Hardy, P., Jouvent, R., & Widlöcher, D. (1984). Speech pause time and the Retardation Rating Scale for Depression (ERD): Towards a reciprocal validation. *Journal of Affective Disorders, 6*, 123-127.

Hargreaves, W. A., & Starkweather, J. A. (1965). *Vocal and verbal indicators of depression.* Unpublished manuscript, University of California Medical School, San Francisco.

Hargreaves, W. A., Starkweather J. A., & Blacker, K. H. (1965). Voice quality in depression. *Journal of Abnormal Psychology, 70*, 218-220.

Heitman, S. (1980). *Sprechstimmveranderungen bei Psychotikern. (Voice changes in psychotic patients)* Unpublished diploma thesis, University of Giessen.

Henze, R. (1954). Experimentelle Untersuchungen zur Phänomenologie der Sprechgeschwindigkeit. (Experimental studies on the phenomenology of speech rate). *Zeitschrift für experimentelle und angewandte Psychologie, 1*, 214-243.

Hinchliffe, M. K., Lancashire, M., & Roberts F. J. (1971). Depression: Defence mechanisms in speech. *British Journal of Psychiatry, 118,* 471-472.

Hoffmann, G. M. A., Gonze, J. C., & Mendlewicz, J. (1985). Speech pause time as a method for the evaluation of psychomotor retardation in depressive illness. *British Journal of Psychiatry, 146,* 535-538.

Hooper, D., Vaughan, P. W., Hinchliffe, M. K., & Roberts, F. J. (1978). The melancholy marriage: An inquiry into the interaction of depression. V. Power. *British Journal of Medical Psychology, 51,* 387-398.

Isserlin, M. (1925). Psychologisch-phonetische Untersuchungen. II. Mitteilung. (Psychological-phonetic studies. 2nd communication). *Zeitschrift für die Gesamte Neurologie und Psychiatrie, 94,* 437-448.

Jones, J. H., & Pansa, M. (1979). Some nonverbal aspects of depression and schizophrenia occurring during the interview. *Journal of Nervous and Mental Disease, 167,* 402-409.

Kasl, S. V., & Mahl, G. F. (1965). The relationship of disturbances and hesitation in spontaneous speech to anxiety. *Journal of Personality and Social Psychology, 1,* 425-433.

Klos, T. (1984). *Psycholinguistische Aspekte im Interview mit depressiven Patienten.* (Paralinguistic aspects in the interview with depressives). Unpublished diploma thesis, University of Giessen.

Klos, T., & Ellgring, H. (1984). Sprechgeschwindigkeit und Sprechpausen von Depressiven. (Speech rate and pauses in the speech of depressives). In M. Hautzinger & R. Straub (Hrsg.), *Psychologische Aspekte depressiver Störungen* (Psychological aspects of depressive disturbances) (pp. 4-26). Regensburg: S. Roderer.

Klos, T., & Ellgring, H. (1987). Manuelle versus elektronische Analyse von Sprechpausen. (Manual vs. electronic measurement of speech pauses). *Zeitschrift für experimentelle und angewandte Psychologie, 34,* 63-71

Kraepelin, E. (1921). *Manic-depressive insanity and paranoia* (Barclay, M., Trans.). Edinburgh: Livingstone.

Kuny, S., & Stassen, H. H. (1993). Speaking behavior and voice sound characteristics in depressive patients during recovery. *Journal of Psychiatric Research, 27,* 289-307.

Ladd, D. R., Scherer, K. R., & Silverman, K. (1986). An integrated approach to studying intonation and attitude. In C. Johns-Lewis (Ed.), *Intonation in discourse.* London/Sydney: Croom Helm.

Levin, S., Knight, R. A., Hall, J. A., & Alpert, M. (1985). Verbal and nonverbal expression of affect in speech of schizophrenic and depressed patients. *Journal of Abnormal Psychology, 94,* 487-497.

Long, G. T. (1988). The relationship of voice stress, anxiety, and depression to life events and personal style variables. *Social Behavior and Personality, 16,* 133-145.

Mahl, G. F. (1965). Disturbances and silences in the patient's speech in psychotherapy. *Journal of Abnormal and Social Psychology, 53,* 1-15.

Mahl, G. F., & Schulze, G. (1964). Psychological research in the extralinguistic area. In T. A. Sebeok, A. S. Hayes, & M. C. Bateson (Eds.), *Approaches to semiotics* (pp. 51-124). The Hague: Mouton.

Mandal, M. K., Srivastava, P., & Singh, S. K. (1990). Paralinguistic characteristics of speech in schizophrenics and depressives. *Journal of Psychiatric Research, 24,* 191-196.

Markel, N. N., Bein, M. F., & Phillis, J. A. (1973). The relationship between words and tone-of-voice. *Language and Speech, 16,* 15-21.

Maser, J. D. (Ed.). (1987). *Depression and expressive behavior.* Hillsdale, NJ: Erlbaum.

Moses, P. (1954). *The voice of neurosis.* New York: Grune & Stratton.

Murray, D. C. (1971). Talk, silence and anxiety. *Psychological Bulletin, 75,* 244-260.

Natale, M. (1977). Effects of induced elation-depression on speech in the initial interview. *Journal of Consulting and Clinical Psychology, 45,* 45-52.

Newman, S. S., & Mather, V. G. (1938). Analysis of spoken language of patients with affective disorders. *American Journal of Psychiatry, 94,* 913-942.

Nilsonne, A. (1987). Acoustic analysis of voice changes during depression and after improvement. *Acta Psychiatrica Scandinavia, 76,* 235-245.

Nilsonne, A., Sundberg, J., Ternström, S., & Askenfelt, A. (1988). Measuring the rate of change of voice fundamental frequency in fluent speech during mental depression. *Journal of the Acoustical Society of America, 83,* 716-728.

Ostwald, P. F. (1964). Acoustic manifestations of emotional disturbances. *Disorders of Communication-Research Publications, 42,* 450-465.

Pittam, J., & Scherer, K. R. (1993). Vocal expression and communication of emotion. In M. Lewis & J. M. Haviland (Eds.), *Handbook of emotions* (pp. 185-198). New York: Guilford.

Pope, B., Blass, T., Siegman, A. W., & Raher, J. (1970). Anxiety and depression in speech. *Journal of Consulting and Clinical Psychology, 35,* 128-133.

Rochester, S. R. (1973). The significance of pauses in spontaneous speech. *Journal of Psycholinguistic Research, 2,* 51-81.

Roessler, R., & Lester, J. (1976). Voice predicts affect during psychotherapy. *Journal of Nervous and Mental Disease, 3,* 166-176.

Sanctis, S. de (1904). *La mimica del pensiero.* (The expression of thought). Milano, Italy: Sandron.

Schatzberg, A. F., Cole, J. O., & Blumer, D. P. (1978). Speech blockage: Tricyclic side effects. *American Journal of Psychiatry, 135,* 600.

Scherer, K. R. (1979). Personality markers in speech. In K. R. Scherer & H. Giles (Eds.), *Social markers in speech* (pp. 147-209). Cambridge, UK: Cambridge University Press.

Scherer, K. R. (1981). Speech and emotional states. In J.K. Darby (Ed.), *Speech evaluation in psychiatry* (pp. 189-220). New York: Grune & Stratton.

Scherer, K. R. (1982). Methods of research on vocal communication: Paradigms and parameters. In K. R. Scherer & P. Ekman (Eds.), *Handbook of methods in nonverbal behavior research* (pp. 136-198). Cambridge, UK: Cambridge University Press.

Scherer, K. R., Ladd, D. R., & Silverman, K. E. A. (1984). Vocal cues to speaker affect: Testing two models. *Journal of the Acoustical Society of America, 76,* 1346-1356.

Scherer, K. R. (1986). Voice, stress, and emotion. In M. H. Appley & R. Trumbull (Eds.), *Dynamics of stress* (pp. 159-181). New York: Plenum.

Scherer, K. R. (1987). Vocal assessment of affective disorders. In J. D. Maser (Ed.), *Depression and expressive behavior* (pp. 57-82). Hillsdale, NJ: Erlbaum.

Scherer, K. R. (1989) Vocal correlates of emotion. In H. Wagner & A. Manstead (Eds.), *Handbook of psychophysiology: Emotion and social behavior* (pp. 165-197). London: Wiley.

Scripture, E. W. (1921). A study of emotions by speech transcription. *Vox, 31,* 179-183.

Siegman, A. W. (1979). Cognition and hesitation in speech. In A. W. Siegman (Ed.), *Of speech and time* (pp. 151-178). Hillsdale, N.J.: Erlbaum.

Siegman, A. W. (1987a). The pacing of speech in depression. In J. D. Maser (Ed.), *Depression and expressive behavior* (pp. 83-102). Hillsdale, NJ: Erlbaum.

Siegman, A. W. (1987b). The telltale voice: Nonverbal messages of verbal communication. In S. Feldstein & A. W. Siegman (Eds.), *Nonverbal behavior and communication,* 2nd ed. (pp. 351-433). Hillsdale, NJ: Erlbaum.

Siegman, A. W., & Feldstein, S. (Eds.) (1979). *Of speech and time.* Hillsdale, NJ: Erlbaum.

Siegman, A. W., & Pope, B. (1972). The effect of ambiguity and anxiety on interviewee verbal behavior. In A.W. Siegman & B. Pope (Eds.), *Studies in dyadic communication* (pp. 29-67). New York: Pergamon.

Smith, P. M. (1979). Sex markers in speech. In K. R. Scherer & H. Giles (Eds.), *Social markers in speech* (pp. 109-145). Cambridge, UK: Cambridge University Press.

Standke, R. (1992). *Methoden der digitalen Sprachverarbeitung in der vokalen Kommunikationsforschung.*(Methods of digital speech analysis in communication research). Frankfurt: Peter Lang.

Stassen, H. H., Bomben, G., & Günther, E. (1991). Speech characteristics in depression. *Psychopathology, 24,* 88-105.

Szabadi, E., Bradshaw, C. M., & Besson, J. A. O. (1976). Elongation of pause time in speech: A simple, objective measure of motor retardation in depression. *British Journal of Psychiatry, 129,* 592-597.

Teasdale, J. D., & Fennel, M. J. V. (1982). Immediate effects on depression of cognitive therapy interventions. *Cognitive Therapy and Research, 6,* 343-351.

Teasdale, J. D., & Fogarty, S. J. (1979). Differential effects of induced mood on retrieval of pleasant and unpleasant events from episodic memory. *Journal of Abnormal Psychology, 88,* 248-257.

Teasdale, J. D., Fogarty, S. J., & Williams, J. M. G. (1980). Speech rate as a measure of short-term variation in depression. *British Journal of Social and Clinical Psychology, 19,* 271-278.

Tolkmitt, F., Helfrich, H., Standke, R., & Scherer, K. R. (1982). Vocal indicators of psychiatric treatment effects in depressives and schizophrenics. *Journal of Communication Disorders, 15,* 209-222.

Tolkmitt, F. J., & Scherer, K. R. (1986). Effects of experimentally induced stress on vocal parameters. *Journal of Experimental Psychology: Human Perception and Performance, 12,* 302-313.

Vanger, P., Summerfield, A. B., Rosen, B. K., & Watson, J. P. (1992). Effect of communication content on speech behavior of depressives. *Comprehensive Psychiatry, 33,* 39-41.

Weintraub, W., & Aronson, H. (1967). The application of verbal behavior analysis to the study of psychological defense mechanisms: IV. Speech pattern associated with depressive behavior. *Journal of Nervous and Mental Disease, 144,* 22-28.

Whitman, E. N., & Flicker, D. J. (1966). A potential new measurement of emotional state: A preliminary report. *Newark Beth-Israel Hospital, 17,* 167-172.

Zerssen, D. von, Köller, D. M., & Rey, E. R. (1970). Die Befindlichkeits-Skala (B-S)—ein einfaches Instrument zur Objektivierung von Befindlichkeitsstörungen, insbesondere im Rahmen von Längsschnittuntersuchungen. (The scale of well-being: A simple instrument to assess mood disorders, particularly in the context of longitudimal studies). *Arzneimittelforschung—Drug Research, 20,* 915-918.

Zuberbier, E. (1957). Zur Schreib- und Sprechmotorik der Depressiven. (On the writing and speaking motor behavior in depressives). *Zeitschrift für Psychotherapie und medizinische Psychologie, 7,* 239-249.