Vocabulary Measures for the Evaluation of Therapy Outcome: Re-Studying Transcripts From the Penn Psychotherapy Project

Michael Hölzer , Erhard Mergenthaler , Dan Pokorny , Horst Kächele & Lester Luborsky

Published online: 25 Nov 2010.

Submit your article to this journal

Article views: 60

View related articles

Citing articles: 5 View citing articles

VOCABULARY MEASURES FOR THE EVALUATION OF THERAPY OUTCOME: RE-STUDYING TRANSCRIPTS FROM THE PENN PSYCHOTHERAPY PROJECT

Michael Hölzer, Erhard Mergenthaler, Dan Pokorny, and Horst Kächele
University of Ulm

Lester Luborsky
University of Pennsylvania

In this study we investigated how vocabularies of patients, as well as therapists, are related to outcome of psychotherapeutic treatment. For this purpose 80 transcripts of the ten most and ten least improved patients of the Penn Psychotherapy Project (PPP) were included in the Ulm Textbank. In addition to the total amount of speech, different vocabulary measures, such as the "private vocabulary" and the "shared vocabulary" of patient and therapists, were applied to the transcripts. Furthermore, the "Regressive Imagery Dictionary" (RID) was applied as a content analytic approach, measuring primary and secondary process content in a given text. Significant correlations of these variables with the PPP outcome measure "residual gain" were found, indicating, that successful therapists tend to accommodate more to the language of their patients compared with their non-successful colleagues.

INTRODUCTION

A variety of computer-aided methods for evaluating psychotherapy process and outcome has been suggested during the last three decades (Lorenz & Cobb, 1954; Spence, 1968; Dahl, 1974; Kächele, 1976; Speidel, 1979; Cierpka, 1980; Wirtz & Kächele, 1983; Kächele, Hohage, & Mergenthaler, 1983; Reynes, Martindale, & Dahl, 1984; Mergenthaler, 1985; Grünzig & Mergenthaler, 1986; Mergenthaler & Pokorny, 1989; Bucci & Mergenthaler, 1993). Although promising in many respects, vocabulary measures are mentioned only rarely in literature on computer-aided strategies in the field of psychotherapy research (Hölzer, Scheytt, Mergenthaler, & Kächele, 1994; de la Parra, 1985). However, systematic studies

on vocabularies or on the change of vocabularies during therapy are not mentioned in the literature at all (Garfield & Bergin, 1986), although it seems to be a straightforward hypothesis that interaction processes like those in psychotherapy should somehow be reflected in language and thus in the vocabularies of the speakers involved.

This almost complete lack of systematic vocabulary studies might be due to the fact that computer supported techniques for identifying linguistic markers as indicators for therapeutic processes face a general dilemma. On the one hand, the rather mechanical way of analyzing data by means of computerized methods paves the way for a host of various procedures; on the other hand, the selection of the variables to be investigated is often restricted to formal criteria lacking clinical relevance.

It was the aim of this study to show that vocabularies and certain features of vocabularies can well be linked to important aspects of process and outcome of psychotherapy. The transcripts from the Penn Psychotherapy Project (Luborsky et al., 1988) seemed to be particularly interesting material, since, on the basis of two outcome measures, patients who benefited from psychotherapy ("most improved") were distinguished from patients who did not benefit ("least improved"). Thus the way was open to investigate how certain features of vocabularies of patients and therapists relate to outcome. As a side effect, a first implementation of English data in the Ulm Textbank system became possible (Hölzer et al., 1994).

Within the Penn Psychotherapy Project (PPP), Luborsky et al. had identified two types of patients' experiences of being in a helpful relationship with a therapist: The Helping Alliance type I and type II, which tended to be observed in therapies with favorable outcomes. Probably the most interesting finding of the PPP was that such an experience accounts for more outcome variance than most of the pretreatment variables examined. Thus it could be concluded that predicting outcome is more successful with in-treatment information about the communication between therapist and patient rather than patient and therapist variables that leave out the actual therapeutic interaction.

The guiding rationale for the work we report here is that successful psychotherapy and the concomitant experience of a "helping alliance" should in some way be reflected in characteristics of both patient and therapist speech, i.e., "vocabulary." Here the term "vocabulary" refers to three aspects of the verbal data of patient and therapist: 1. "Verbal activity" measured by the total amount of text (= token) produced by a single speaker or the dyad. 2. The size of the "private" as well as the "shared" vocabularies of patient and therapist. 3. The amount of material that can be considered to be characteristic for primary process as well as secondary process thinking, measured by the Regressive Imagery Dictionary (Martindale, 1975) a classic among computerized content analytic procedures.

MATERIAL

Before the methods and findings are presented, a brief description of the PPP sample examined will be given: The Penn Project has led to the identification of the ten most and ten least improved patients who were treated for at least 25

sessions and selected on the basis of two outcome measures "residual gain" and "rated benefits" from a total sample of 73 cases.

"The two outcome measures, rated benefits and residual gain, were devised as follows: First, composite measures of pretreatment and posttreatment adjustment were compiled using ratings made by a clinical observer and the patient (weighted equally). The measures included the Inventory of Social and Psychological Functioning (A. Auerbach, M. Johnson, unpublished data, December 1977); Minnesota Multiphasic Personality Inventory scales for ego strength, hypochondriasis, and hysteria; the health sickness Rating Scale = HSRS (Luborsky, 1962; Luborsky, 1975), and the adjustment items of the Prognostic Index Interview (Auerbach, Luborsky, & Johnson, 1972). Second, a measure of raw gain was obtained by subtracting pretreatment scores from posttreatment scores. Third, by partialling out the pretreatment scores from raw gain, a residual gain score was obtained that showed the gain of each patient relative to all other patients as if each one had begun treatment at the same level. Finally, Rated Benefits scores were obtained by getting direct ratings of change from patients and therapists independently. Residual Gain and Rated Benefits scores were highly correlated (.76). Seven of the most improved and eight of the least improved patients were chosen for study in terms of Residual Gain scores and the additional five of the sample of 20 were chosen in terms of Rated Benefits scores. Since the two criteria were highly correlated, cases were taken that were extreme on either criterion" (Morgan, Luborsky, Crits-Christoph, Curtis, & Solomon, 1982, p. 398).

The 20 patients were treated by 18 psychodynamically oriented, male therapists, ten of them second or third-year residents under supervision and eight of them more experienced. Since "age" and "experience" of a therapist sometimes shape attitudes of a therapist more than her or his original theoretical orientation, we used them both as independent variables possibly influencing the verbal behavior of therapist as well as patient.

There were no significant differences between the most versus the least improved cases with respect to their demographic patient characteristics. According to *DSM-II* all 20 were nonpsychotic outpatients, although they differed considerably in their diagnoses (e.g., adjustment reaction of adolescence, obsessive-compulsive personality disorder, depressive reaction, and schizoid personality). "Age of patient" (ranging from 21 to 47 with the most improved, from 18 to 39 with the least improved) was included as an additional variable in our study.

The median length of treatment was 61 weeks for the most improved, and 43 weeks for the least improved. The therapy sessions were audiotaped and transcribed. For each patient four segments were drawn, two from the initial (sessions 3 and 5) as well as two from the final stage of treatment, "the session at which 90 percent of the treatment had been completed and the prior session (Morgan et al., 1982). Thus, a total of 80 segments resulted. Each segment consisted of the first 20 minutes of a session. In order to facilitate the evaluation, the early-in-treatment segments were combined, as were the late-in-treatment segments, yielding two "texts" for each patient, one for each stage of therapy.

METHOD

"Residual gain" as an outcome criterion was used both as grouping variable for distinguishing "the most improved" from "the least improved" and (together with "age of patient," "age of therapist" as well as "experience of therapist") as independent variable for correlation with the three aspects of vocabularies mentioned above: 1. Verbal activity; 2. Vocabulary size; and 3. Scores of Regressive Imagery Dictionary categories as a content analytic measure.

VERBAL ACTIVITY

In the third edition of the *Handbook of Psychotherapy and Behavior Change* (Garfield & Bergin, 1986), Orlinsky and Howard state that "if the patient doesn't talk much of the time and the therapist doesn't talk some of the time, then it would be hard to claim that psychotherapy is actually taking place" (Orlinsky & Howard, 1986, p. 320). In spite of its objectivity and the ease with which it is obtained, verbal activity defined as the total number of words (token) occurring in a given text is often seen as too simple a measure to reflect the complexity of what is going on in psychotherapy and thus viewed as theoretically unimportant. However, this attitude among psychotherapy researchers might reflect their prejudice against simplicity, rather than empirically supported facts. There is still not enough well established knowledge about how the amount of patient's and therapist's speech relate to each other and how that influences psychotherapeutic processes. Orlinsky and Howard point out that a considerable number of studies have shown that verbal activity is significantly linked to outcome: "A majority of findings (7 of 11) relating the amount of patient speech to outcome were significantly positive" (1986, p. 320). In studies of therapist speech, only 5 of 14 showed significantly positive findings "while 8 showed no relationship of therapist talkativeness to patient benefit." However, when studying the phenomena in detail, wide variations of verbal activity may be observed especially in longer treatments (Kächele et al., 1983). Furthermore, the "balancing" of patient and therapist conversational activity, measured by the ratio of therapist to patient verbal activity, might be a more potent predictor of outcome than simple activity measures (Scobel, 1979; Zimmer & Cowles, 1972). Since in some cases a simple ratio of therapist to patient speech (or visa versa) can be highly misleading for a variety of mathematical reasons, we computed ratios of verbal activities with more appropriate mathematical characteristics like for example: the-va/pat-va = 100 × (the-va/the-va + pat-va), where "the-va" is the verbal activity of the therapist and "pat-va" the verbal activity of the patient.

Hypothesis. Since Scobel reports a 1 to 2 ratio in therapist-verbal-activity to patient-verbal-activity for successful and a 1 to 3 ratio for non-successful Rogerian psychotherapy, we expected to find similar results.

VOCABULARY MEASURES

As opposed to verbal activity, formal vocabulary measures do not belong to the present battery of psychotherapeutic research tools although they might well help fill the gap between formal and content related approaches. As opposed to verbal activity the term "vocabulary" refers to the number of different words (types) that are used by a speaker. The ratio between types and token, the TTR,

has been sometimes looked at as an indicator of the diversity of a text. From a research perspective, vocabulary measures defined in terms of types are interesting, not only because they are easily and objectively obtained. Since words stand for concepts (and therapy has essentially to do with an exchange of concepts and beliefs, with assimilation of new material and accommodation of previous schemata), changes in vocabularies during treatment might parallel or at least partly reflect such learning processes.

In a therapeutic dialogue, different kinds of vocabularies can be distinguished (de la Parra, 1984). For matter of presentation we focus on two of them studied in this investigation:

1. The "Private Vocabulary" (PV), i.e., the set of words or types that are used by only one of the speakers,[1] here denoted as Patient-PV and Therapist-PV.
2. The "Shared Vocabulary" (SV), the set of types that are used by both patient and therapist, again denoted as "Patient-SV" or "Therapist-SV."

Hypothesis. The ability of a therapist to accommodate to the language of his or her patient, to bridge social differences and to empathize with the patient should result in a low "Therapist-PV" on his part. Therapists of the most improved patients are likely to show more of these capacities than therapists of the least improved, i.e. show smaller Private Vocabularies.

THE REGRESSIVE IMAGERY DICTIONARY

As a third method we applied the Regressive Imagery Dictionary (RID) (see Figure 1) as developed by Martindale (1975) for use in computerized content analysis to the transcripts. Psychoanalytic theory holds that in order for change to occur in the patient, regression toward primary process thinking must be achieved. Thus, free association, the "basic rule" of psychoanalysis, is generally viewed as inducing such a movement from secondary process cognition to the more primitive and less mature primary process. Although in the light of recent findings of cognitive psychology the theory of the "regression in the service of the ego" (Kris, 1952, p. 370) needs some revision, the categories with which Martindale's Regressive Imagery Dictionary measures primary and secondary process phenomena are nevertheless interesting.

"Primary process is based upon the occurrence of words with connotations of Drives, Sensations, Defensive Symbolization (i.e., references to disorder in the external world), Regressive Cognition (direct references to alteration in states of consciousness), and Icarian Imagery (i.e., references to fire, water, rising and falling). Secondary process tabulates references to abstraction, instrumental behavior, social behavior, time, moral imperatives, order and restraint" (Reynes et al., 1984, p. 734). In its original version the RID contains a total of about 2900

---

[1] A slightly more sophisticated way to compute the Private Vocabulary results in what we call the "Characteristic Vocabulary." Here the decision as to whether a certain type belongs to the vocabulary is based on probabilities rather than simple "yes or no" decisions. A word has to occur in the text of one speaker a number of times as often as in the text of another to be incorporated in his Characteristic Vocabulary. Depending on the chosen probability, the magnitude of the Characteristic Vocabulary may differ considerably.

Primary Process

Drives
01 orality
02 anality
03 sex

Sensation
04 general sensation
05 touch
06 taste
07 odor
08 sound
09 vision
10 cold
11 hard
12 soft

Defensive Symbol.
13 passivity
14 voyage
15 random movement
16 diffusion
17 chaos

Regressive Cognition
18 unknown
19 timelessness
20 concious alter
21 brink passage
22 narcissism
23 concreteness

Icarian Imagery
24 ascend
25 height
26 descend
27 depth
28 fire
29 water

Secondary Process

Abstract Thought
30 abstract thought

Behavior
31 social behavior
32 instrumental behavior
33 restraint
34 order
35 temporal reference
36 moral imperative

Emotions

37 positive affects
38 anxiety
39 sadness
40 affection
41 aggression
42 expressive behavior
43 glory

Figure 1. Categories and subcategories of the "Regressive Imagery Dictionary."

words, which are assigned to 43 exclusive categories. The Ulm version of the RID works with a total number of more than 5000 words, reduced to their basic forms.

Categories 1 to 29 make up the primary process, while categories 30 to 36, yield the secondary process. A third subset of categories (37 to 43), labeled "emotions," were also assessed in the transcripts. They contain words referring to positive and negative affects such as anxiety, sadness, anger, and so on.

The Regressive Imagery Dictionary has shown good construct validity in studies of a wide variety of texts. As Reynes, Martindale, and Dahl (1984) report: "More primary process and less secondary process imagery have been found in folk tales of primitive as opposed to complex preliterate societies, in poetry of writers who exhibit signs of psychopathology as opposed to writers who do not show such signs and in stories told by younger as opposed to older children" (p. 734).

Evaluating working, neutral, and resistance sessions of a psychoanalytic therapy by means of the RID, these authors found that primary process increases and secondary process decreases significantly as one moves from resistance through neutral to working sessions. Because patient and therapist "did not exhibit radically different amounts of primary process in their speech," they concluded that "psychoanalytic work involves a movement toward primary process cognition on the part of both patient and analyst." (p. 736).

Hypothesis. Although the PPP patients in the present study were treated by psychoanalytically informed psychotherapy and not psychoanalysis proper, we

hypothesized that success in supportive-expressive dynamic therapy should also positively correlate with the primary process and the emotion subset of categories as well.

DATA ANALYSES

The main analyses were *t*-tests for independent groups to test for significant differences between the two outcome groups (i.e., more or less improved patients). To test for significant changes during treatment within each group (i.e., early-in treatment versus late-in treatment sessions) paired *t*-tests were applied. Pearson *r* was calculated for correlations between independent variables (like age of patient and therapist as well as experience of therapist) and dependent variables (i.e., verbal activity, the size of Private as well as Shared Vocabularies, and the categories of the Regressive Imagery Dictionary).

RESULTS

VERBAL ACTIVITY

For both groups the mean ratio of therapist talkativeness to patient talkativeness was 1/3.5, meaning that patients in general talked three times as much as their therapists. Thus, "psychotherapy took place" since patients talked "much" of the time and therapists "some of the time." However, both groups showed considerable within-groups variability. Patients (as well as therapists) did not exhibit any differences in their overall verbal activity, that is, in their average scores over all 4 sessions, between the groups. In this respect their conversational behavior seemed to be almost identical. Thus, we could not confirm Scobel's (1979) results. Still, an interesting difference between the most improved and the least improved was found when we looked at this ratio at the different points of time in treatment (see Figure 2).

Moving from the initial to the final stage in therapy the ratio drops with the most improved and rises with the least improved, meaning that therapists of the most improved talk relatively less and therapists of the least improved talk relatively more in relationship to their patients late in treatment compared with the beginning. Perhaps these therapists needed to say less for the most improved patients because they were proceeding at an adequate pace; the therapists of least improved patients needed to say more because the treatment was not going as well as it should. But, almost significant at a 5 percent level, this interaction effect of "time" and "outcome" with respect to the ratio of verbal activity renders simple verbal activity measures almost meaningless and has to be further examined.

Consistent with this, we found a highly significant negative correlation between patients' verbal activity and "Residual Gain" with non-improvers. (See Table 1.)

Whereas a verbally active therapist tends to have a negative impact on improving patients (*r* = -.50), more talkative patients among the non-improvers show clearly the least improvement of this group.
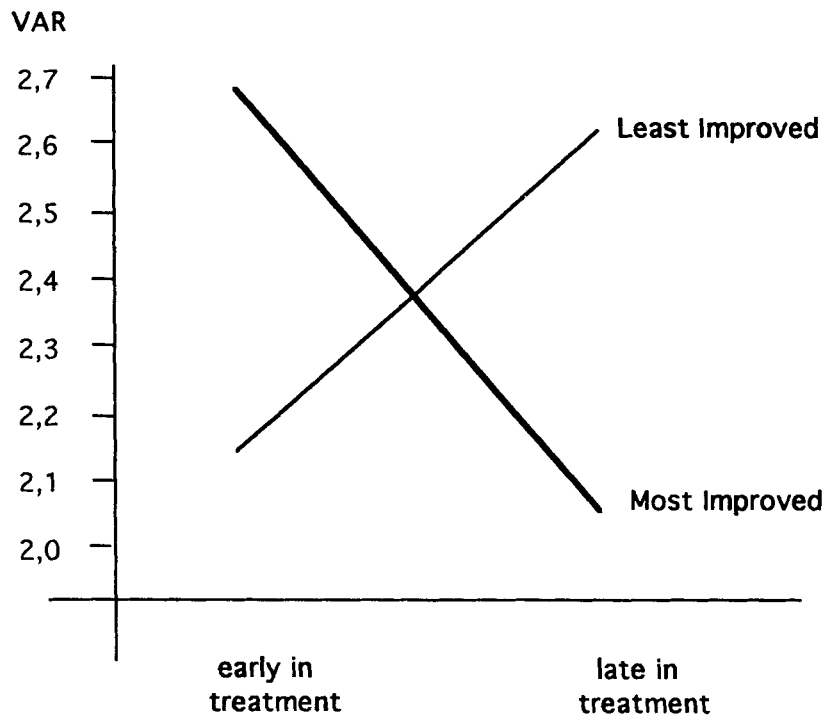
VAR

Figure 2. The change in verbal activity ratio = VAR [100 × (Token Th/Token Th + Token Pat)] throughout therapy. For the interaction effect between "time" and "group" *t* was —2.01 (*t*-test for independent groups; *p* = 0.059).

PRIVATE AND SHARED VOCABULARIES

Although not as impressive as in verbal activity, a considerable variation in their Private (PV) and Shared Vocabularies (SV) is seen in both therapy groups. Comparing the vocabulary measures between the most improved and the least improved does not lead to significant differences, with one exception.

The correlation between residual gain as outcome measure and the size of Patient-PV distinguishes improvers (they show a mildly positive correlation) and non-improvers, with whom we found a highly significantly negative correlation. Again, this indicates an interaction effect. The Private Vocabulary of improved patients might well reflect personal attitudes and qualities of the patient, such as

Table 1. Correlation Coefficients (Pearson *r*) between "Verbal Activity" and "Residual Gain"

	Therapist	Patient
Most improved	-.50	.23
Least improved	.19	.79**

Verbal Activity was averaged over all 4 sessions. (*N* = 10 in each group).
** = *p* < .01.

Table 2. Correlation Coefficients (Pearson r) between "Private Vocabularies" and "Residual Gain" as well as "Age" and "Experience" of Therapists

	Private Vocabularies	
	Private Voc. Therapist	Private Voc. Patient
Most improved		
Res. Gain	-.59*	.44
Age (The)	-.36	—
Exp. (The)	-.33	—
Least improved		
Res. Gain	—	-.83**
Age (The)	-.38	—
Exp. (The)	-.59*	—

$N = 10$ in each group.
* $= p < .10$.
** $= p < 0.01$.

relative autonomy or certain cognitive capacities, whereas a large Private Vocabulary with non-improved patients may be a hint that there is a rather monologic style of conversation, the content of which is not being fully processed and integrated by the therapist (see Table 2). We could not support our hypothesis of a lower Therapist-PV among improvers. Furthermore, in both groups the Private Vocabularies stay essentially the same throughout the treatment; there are no differences discernable between the initial and the end phases of therapy.

Although not significant, the negative correlations between the therapist's experience and her or his Private Vocabulary, which was found in both groups, seem to be of interest. Here perhaps it is fair to speculate that experienced therapists tend not to distance themselves from the lexical capacities of their patients.

The Shared Vocabulary between patient and therapist exhibits a significant decrease with the most improved patients: From a mean of 241 for the early-in-treatment text it drops to 207 words for the late-in-treatment segments ($t = 2.32$, $p < .05$). Obviously, this finding runs counter to our expectation of the reverse relationship, viewing an increase in the Shared Vocabulary as an indicator for a more successful therapeutic exchange. Here, our knowledge about the Shared Vocabulary is not yet sufficient enough to formulate valid hypotheses. This vocabulary seems to be more complicated and composed of a variety of sub-vocabularies, that is, words that are necessary to construct meaningful language as well as words that refer to specific and maybe therapeutically important content categories. Our next step will be to determine these sub-vocabularies and to formulate differential hypotheses.[2]

REGRESSIVE IMAGERY DICTIONARY (RID)

While 84.2 percent of the text (of the 80 transcripts) did not fall into one of the categories of the Regressive Imagery Dictionary, 2.3 percent were assigned to the emotion category, 2.1 percent to the Primary Process categories and 11.4

[2]A closer examination of the "Characteristic Vocabularies" yielded the finding that successful therapists (that is, therapists of "the most improved patients") tend to focus particularly on the emotional parts of the patient's vocabulary.

percent to the Secondary Process categories. Compared with findings of other studies (where this percentage is varying from about 3 to 7 percent), an average of 2.1 percent primary process of the total text is a rather low percentage. The 1:5 ratio of primary to secondary process found in this study compares best with the resistance hours identified by Reynes et al. (1984).

Neither patients nor therapists exhibited significantly different amounts of primary and secondary process in their speech, although there is a tendency of therapists having higher scores in primary process than patients. This is true for both improvement groups, which do not differ from each other in the RID categories either.

The concept "regressive cognition" as well as the category "sensation" covering single categories like touch, taste, odor, sound or vision contribute significantly more to the primary process score than the other concepts, both in patient and in therapist speech. (See Figure 3.)
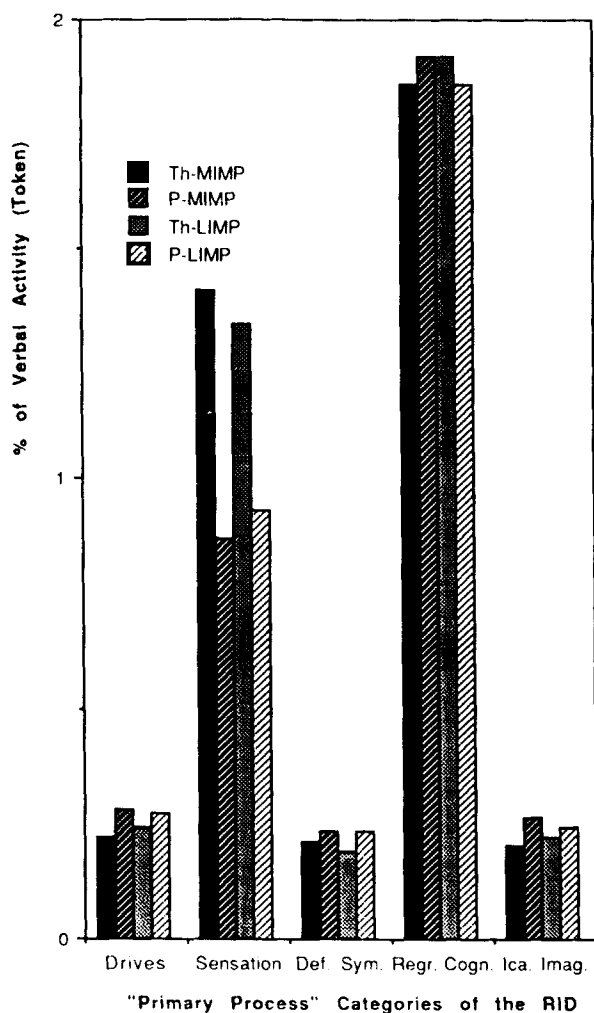

Figure 3. Primary process categories in percentage of the Verbal Activity (sum scores of 4 sessions, Th = therapist, P = patient, MIMP = most improved, LIMP = least improved).

Furthermore, in both groups, therapists utter more words listed in the sensation category than their patients. This finding was highly significant. Although the sensation category does not discriminate poor from successful outcome, this result may hint at a general therapeutic attitude or stance. It can be interpreted in the light of the Dual Code Theory and Bucci's findings about referential activity (Bucci, 1985), which refer to two different representational systems operating in the mind, a verbal and a nonverbal one. In this theory, referential activities are seen as linking or bridging functions between the two systems and manifested in concrete and specific language and by the use of words which refer to perceptual or sensory input. Maybe therapists, by preferring a rather concrete way of talking and thus revealing high levels of referential activity themselves, try to induce referential activity on the part of their patients. In a recent study this hypothesis was tested by correlating the RID measures with ratings of referential activity (Schnee, 1994).

No changes at all took place throughout therapy in terms of primary process categories, except that improved patients show a significant rise in "icarian imagery" (paired t-test, $t = 2,54$; $p < .05$). However, in both groups patients show a distinct drop in secondary process, mostly due to a substantial decrease of "abstract thought" (significant at the 10 percent level in each group; improved patients, paired t-test: $t = -1.97$, $p < .10$; non-improved patients, paired t-test: $t = -1.84$, $p < .10$) whereas the second secondary process category "behavior" does not change at all. The drop in "abstract thought" becomes distinctly significant, when all patients are looked at as one group (mean drop: -.58%, $t = -2,76$; $p < 0.02$). Therapists of both groups tend to have lower levels of secondary process in the late-in-treatment sessions. Here, only the drop in "behavior" for therapists of improved patients turned out to be significant (paired t-test: $t = -2.28$, $p < .05$). Thus, one might claim that, although the primary process levels in terms of absolute scores stay the same for all speakers throughout therapy, the ratio between primary and secondary process rises due to the drop in the secondary process categories. (See Table 3.)

DISCUSSION

The following conclusions can be drawn from our findings:

1. It is obvious that the investigation of just the particles of speech by computer-aided text analysis limits the range of possible conclusions, since investigation of lexical items still is inevitably tied to a neglect of contextual information. However, it takes advantage of a technology which may balance the poverty of the semantics involved (Mergenthaler, 1985).
2. Undoubtedly the linguistic analysis of the Penn transcripts is made more difficult by the wide variety of diagnoses found in the 20 cases examined. Clearly, improvement and the way in which it is reflected linguistically in a case of "latent schizophrenia" is different from improvement in a case of adjustment reaction. Therefore, the measures mentioned above were not expected to fully explain the "linguistic variance" inherent in our sample.

Table 3.   Pearson r Correlations between RID Categories and "Age" of Patient, as well as "Experience" of Therapists and "Residual Gain"

	Most Improved					Least Improved			
Patient	Pat. Age	Therap. Age	Therap. Exp.	Resid. Gain	Patient	Pat. Age	Therap. Age	Therap. Exp.	Resid. Gain
Primary P.	.53				Primary P.				
Secon. P					Second. P				.74**
Emotion	-.48				Emotion				
Drive					Drive		-.61*	-.64*	
Sensation			.54		Sensation				.49
Symbol.	.66**				Symbol.	-.87***			-.70**
Cognition	.51				Cognition				
Icarian	.58	.47	.45		Icarian		.54		
Abstract	-.51	-.56*			Abstract				.56*
Behavior	.86***	0.70**	.51		Behavior	.47			
Therapist					Therapist				
Primary P.				.55*	Primary P.				
Secon. P	.41	.55*	.43		Secon. P				
Emotion					Emotion		-.55*	-.70**	
Drive	-.61*	-.86***	-.66**		Drive				
Sensaton					Sensation			-.53	.46
Symbol.				-.49	Symbol.				
Cognition				-.60*	Cognition				
Icarian	-.53			-.68**	Icarian				
Abstract					Abstract				
Behavior	.44			.68**	Behavior		-.68**	-.56*	

$* = p < .10.$
$** = p < .05.$
$*** = p < .01.$

3. Refinement of the methods applied seems to be necessary on two levels:
   a. The vocabulary methods clearly need revision with respect to the existence of differently relevant sub-vocabularies and related differential hypotheses. Here the investigation of the emotional vocabulary by means of an English version of the "Affective Dictionary Ulm" (Hölzer, Scheytt, & Kächele, 1992) seems to yield particularly interesting results.
   b. Content analytic methods, such as the RID have to be further validated, and computer aided measures for other concepts in psychotherapeutic or psychoanalytic theory have to be developed. Among others, the work of Bucci and Mergenthaler (1993) can be mentioned here, for they have already been able to measure referential activity by means of a specially adjusted dictionary.

In conclusion, it can be said that although the evaluation of psychotherapy transcripts like the Penn material by means of computer aided methods is still in an early stage, we are optimistic that these methods will continue to contribute to the nomological net with which psychotherapeutic processes become more and more explainable.

REFERENCES

Auerbach, A. H., Luborsky, L., & Johnson, M. (1972). Clinicians' predictions of outcome of psychotherapy: A trial of a prognostic index. *American Journal of Psychiatry, 128,* 830-835.

Bucci, W. (1985). Dual coding: a cognitive model for psychoanalytic research. *Journal of the American Psychoanalytic Association, 33,* 571-607.

Bucci, W., & Mergenthaler, E. (1993). Computer assisted procedures for modelling Referential Activity. In S. f. P. Research (Ed.), *1993 Annual International Meeting,* Pittsburgh.

Cierpka, M. (1980). Personalpronomina als Indikatoren für interpersonale Beziehungen in einer psychoanalytischen Gruppen-therapie. *Psychotherapie und Medizinische Psychologie, 30,* 212-217.

Dahl, H. (1974). The measurement of meaning in psychoanalysis by computer analysis of verbal context. *Journal of the American Psychoanalytic Association, 22,* 37-57.

de la Parra, G. (1985). Differentielle Textmaße. Auswertungen an einem psychoanalytischen Erstinterviewkorpus. Ulm: PSZ Verlag.

Garfield, S., & Bergin, A. E. (Eds.) (1986). *Handbook of psychotherapy and behavior change,* (3rd ed.). New York: John Wiley.

Grünzig, H. J., & Mergenthaler, E. (1986). Computerunterstützte Ansätze. Empirische Untersuchungen am Beispiel der Angstthemen. In U. Koch & G. Schöfer (Eds.), *Sprachinhaltsanalyse in psychosomatischen und psychiatrischen Forschung* (pp. 203-212). Weinheim München: Psychologieverlagsunion.

Hölzer, M., Aeschelmann, D., Schwilk, C., & Kächele, H. (1993). Defense mechanisms and linguistic styles. *New Trends in Experimental and Clinical Psychiatry, 9,* 15-22.

Hölzer, M., Scheytt, N., & Kächele, H. (1992). Das "Affektive Diktionär Ulm" als eine Methode D der quantitativen Vokabularbestimmung. In C. Z. a. P. Mohler (Ed.), *Textanalyse. Anwendungen der computerunterstötzten Inhaltsanalyse. Beiträge zur 1. TEXTPACK-Anwenderkonferenz* (pp. 131-154). Opladen: Westdeutscher Verlag GmbH.

Hölzer, M., Scheytt, N., Mergenthaler, M., & Kächele, H. (1994). Der Einfluß des Settings auf die therapeutische Verbalisierung von Affekten. *Psychotherapie Medizinische Psychologie Psychosomatik.*

Kächele, H. (1976). *Maschinelle Inhaltsanalyse in der psychoanalytischen Prozeßforschung.* Ulm: PSZ- Verlag.

Kächele, H., Hohage, R., & Mergenthaler, E. (1983). Therapie-orientierte Dokumentation in einer psychotherapeutischen Ambulanz—Funktion und Implikation. *Psychotherapie Medizinische Psychologie, 33,* 142-146.

Kris, E. (1952). *Psychoanalytic Explorations in Art.* New York: IUP.

Lorenz, M., & Cobb, S. (1954). Language patterns in psychotic and psychoneurotic subjects. *Archives of Neurology and Psychiatry, 72,* 665-673.

Luborsky, L. (1962). Clinicians' judgments of mental health: A proposed scale. *Archives of General Psychiatry, 7,* 407-417.

Luborsky, L. (1975). Clinician's judgments of mental health: specimen case descriptions and forms for the Health-Sickness Rating Scale. *Bulletin of the Menninger Clinic, 35,* 448-480.

Luborsky, L., Crits-Christoph, P., Mintz, J., & Auerbach, A. (1988). *Who will benefit from psychotherapy? Predicting therapeutic outcomes.* New York: Basic Books.

Luborsky, L., Mintz, J., Auerbach, A., Christoph, P., Bachrach, H., Todd, T., Johnson, M., Cohen, M., O'Brien, C. P. (1980). Predicting the outcomes of psychotherapy. Findings of the Penn Psychotherapy project. *Archives of General Psychiatry, 37,* 471-481.

Martindale, C. (1975). *Romantic progression. The psychology of literary history.* Washington, DC: Hemisphere.

Mergenthaler, E. (1985). *Textbank systems. Computer science applied in the field of psychoanalysis.* Berlin/Heidelberg/New York: Springer.

Mergenthaler, E., & Pokorny, D. (1989). Die Wortarten- Verteilung-eine linguo-statistische Textanalyse. In F. Faulbaum, R. Haux, & H. Jöckel (Eds.), *SoftStat 89. Fortschritte der Statistik-Software 2* (pp. 512-521). Stuttgart.: Fischer.

Morgan, R., Luborsky, L., Crits-Christoph, P., Curtis, H., & Solomon, J. (1982). Predicting the outcomes of psychotherapy by the Penn Helping Alliance method. *Archives of General Psychiatry, 39,* 397-402.

Orlinsky, D. E., & Howard, K. I. (1986). Process and outcome in psychotherapy. In S. L. Garfield & A. E. Bergin (Eds.), *Handbook of psychotherapy and behavior change* (pp. 270-318), New York: John Wiley.

Reynes, R., Martindale, C., & Dahl, H. (1984). Lexical differences between working and resistance sessions in psychoanalysis. *Journal of Clinical Psychology, 40,* 733-737.

Schnee, F-P. (1994). Über den Zusammenhang zwischen Referentieller Aktivität und dem Affektiven Vokabular in psychotherapeutischen Erstinterviews. Unpublished doctor thesis, University of Ulm.

Scobel, W. A. (1979). Sprache und Gesprächspsychotherapie. Eine inhaltsanalytische Unter-

suchung. *Psychotherapie Medizinische Psychologie und Psychosomatik, 29,* 66–72.

Speidel, H. (1979). *Entwicklung und Validierung eines Wörterbuches zur maschinell-inhaltsanalytischen Erfassung psychoanalytischer Angst-themen.* Unpublished thesis, Universität Ulm.

Spence, D. P. (1968). The processing of meaning in psychotherapy: Some links with psycholin-

guistics and information theory. *Behavioral Sciences, 13,* 349–361.

Wirtz, E. M., & Kächele, H. (1983). Emotive aspects of therapeutic language: A pilot study on verb-adjective-ratio. In W. M. a. W. Herff (Eds.), *Methodology in psychotherapy research* (pp. 130–135). Frankfurt am Main: Lang.

Zimmer, J. M., & Cowles, K. H. (1972). Content analysis using FORTRAN. *Journal of Counseling Psychology, 19,* 161–166.