

# »Reveal $\Delta_x$ «!

## Zu früheren Projekten der automatisierten Diskursanalyse

Bernhard Dotzler

**Eine Hypothese ist nur unter der Bedingung tauglich,  
daß sie das beabsichtigte Ziel verfehlt,  
um ein anderes, unbekanntes  
zu erreichen.  
Siegfried Kracauer, *Ginster***

**Weimar, 25.10.2018**

BERNHARD  
DOTZLER

ZU FRÜHEREN  
PROJEKTEN  
DER  
AUTOMATI-  
SIERTEN  
DISKURS-  
ANALYSE

DOCUMENT RESUME

ED 041 610

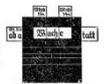
LI 002 069

AUTHOR Stevens, Mary Elizabeth  
TITLE Automatic Indexing: A State-of-the-Art Report.  
INSTITUTION National Bureau of Standards (DOC), Washington, D.C.  
REPORT NO NBS-Monogr-91  
PUB DATE Feb 70  
NOTE 298p.  
AVAILABLE FROM Superintendent of Documents, U.S. Government  
Printing Office, Washington, D.C. 20402 (GPO  
C13.48:91, \$2.25)

ALGORITHMISCHE

DISKURSANALYSE

FAKULTÄT MEDIEN, BAUHAUS-UNIVERSITÄT WEIMAR



BERNHARD  
DOTZLER

ZU FRÜHEREN  
PROJEKTEN  
DER  
AUTOMATI-  
SIERTEN  
DISKURS-  
ANALYSE

Share 1

# WordSmith Tools

Purchase Support Research Downloads

Windows software for finding word patterns

Published by [Lexical Analysis Software](#) and [Oxford University Press](#) since 1996

22nd  
Year!

ALGORITHMISCHE

DISKURSANALYSE



## Concord

... for finding all instances of a word or phrase.

## KeyWords

... helps find salient words in a text or set of texts.

## WordList

... lists the words in your text(s) in alphabetical and frequency order.

and a number of further [Utility tools](#)

or words to that effect.  
or words to ditto ditto  
or words spoken dire

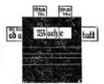
you could play on words keep using  
essential focus on words and the word  
given five sets of words, and asked  
the beginnings of words, and the, the  
about the kind of words you use. No  
passage or set of words is assigned  
big dictionary of words and synonym  
ggins's choice of words in defeat. H  
low what kind of words do you think  
primarily men of words." To put it n  
olicitor a form of words to cover you

fragment of concordance

**The style of a discourse is the message carried by frequency-distributions and transitional probabilities of its linguistic features, especially as they differ from those of the same features in the language as a whole.**

**B. Bloch, Linguistic structure and linguistic analysis (1953),**

**zit. n. Werner Müller, Textklassifikation und Stilanalyse.  
Gedanken zur automatischen Beschreibung eines  
Produktes und seines Produktionsprozesses (1972)**

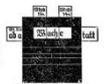


## A Mathematical Theory of Communication

By C. E. SHANNON

6. Second-order word approximation. The word transition probabilities are correct but no further structure is included.

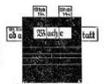
THE HEAD AND IN FRONTAL ATTACK ON AN ENGLISH WRITER THAT THE CHARACTER OF THIS POINT IS THEREFORE ANOTHER METHOD FOR THE LETTERS THAT THE TIME OF WHO EVER TOLD THE PROBLEM FOR AN UNEXPECTED.



**Darling Sweetheart,  
You are my fellow feeling. My affection curiously  
clings to your passionate wish. My liking yearns  
to your heart. You are my wistful sympathy: my  
tender linking.**

**Yours beautifully,  
M.U.C.**

**M.U.C. = Manchester University Computer**



BERNHARD  
DOTZLER

ZU FRÜHEREN  
PROJEKTEN  
DER  
AUTOMATI-  
SIERTEN  
DISKURS-  
ANALYSE

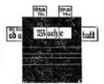
**KEIN KUSS IST STILL  
ODER DIE LIEBE IST STILL  
ODER KEINE SEELE IST REIN  
UND NICHT JEDER KUSS IST GRUEN  
UND EIN JUENGLING IST HEFTIG**

**Rul Gunzenhäuser  
& programmgesteuerte  
Ziffernrechenanlage**

ALGORITHMISCHE

DISKURSANALYSE

FAKULTÄT MEDIEN, BAUHAUS-UNIVERSITÄT WEIMAR



BERNHARD  
DOTZLER

ZU FRÜHEREN  
PROJEKTEN  
DER  
AUTOMATI-  
SIERTEN  
DISKURS-  
ANALYSE

# *The Analysis of Communication Content*

Developments in Scientific Theories  
and Computer Techniques

Edited by  
GEORGE GERBNER  
OLE R. HOLSTI  
KLAUS KRIPPENDORFF  
WILLIAM J. PAISLEY  
PHILIP J. STONE

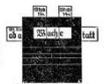
JOHN WILEY & SONS, INC. New York · London · Sydney · Toronto

1969

ALGORITHMISCHE

DISKURSANALYSE

FAKULTÄT MEDIEN, BAUHAUS-UNIVERSITÄT WEIMAR



BERNHARD  
DOTZLER

ZU FRÜHEREN  
PROJEKTEN  
DER  
AUTOMATI-  
SIERTEN  
DISKURS-  
ANALYSE

THE GENERAL INQUIRER:  
A Computer Approach to Content Analysis

*Philip J. Stone a.o.  
Dexter C. Dunphy  
Marshall S. Smith  
Daniel M. Ogilvie  
with associates*



(1966)

The M.I.T. Press  
Massachusetts Institute of Technology  
Cambridge, Massachusetts, and London, England

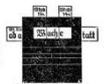
712521

1966

ALGORITHMISCHE

DISKURSANALYSE

FAKULTÄT MEDIEN, BAUHAUS-UNIVERSITÄT WEIMAR



BERNHARD  
DOTZLER

ZU FRÜHEREN  
PROJEKTEN  
DER  
AUTOMATI-  
SIERTEN  
DISKURS-  
ANALYSE

THE GENERAL INQUIRER:  
A Computer Approach to Content Analysis

*Philip J. Stone a.o.  
Dexter C. Dunphy  
Marshall S. Smith  
Daniel M. Ogilvie  
with associates*

**Welcome to the General Inquirer Home Page.**

We mourn the loss of Philip Stone who died on January 31st, 2006, and we dedicate the continuation of this site to his memory. Established as the home website for the General Inquirer, a computer-assisted approach for content analyses of textual data, the site is designed to be a resource for learning about the Inquirer as well as a reference in using the Inquirer.

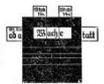
719E

<http://www.wjh.harvard.edu/~inquirer/Home.html>

ALGORITHMISCHE

DISKURSANALYSE

FAKULTÄT MEDIEN, BAUHAUS-UNIVERSITÄT WEIMAR



**... set of computer programs to (a) identify systematically, within texts, instances of words and phrases that belong to categories specified by the investigator; (b) count occurrences and specified co-occurrences of these categories; (c) print and graph tabulations; (d) perform statistical tests; and (e) sort and regroup sentences according to whether they contain instances of a particular category or combination of categories.**



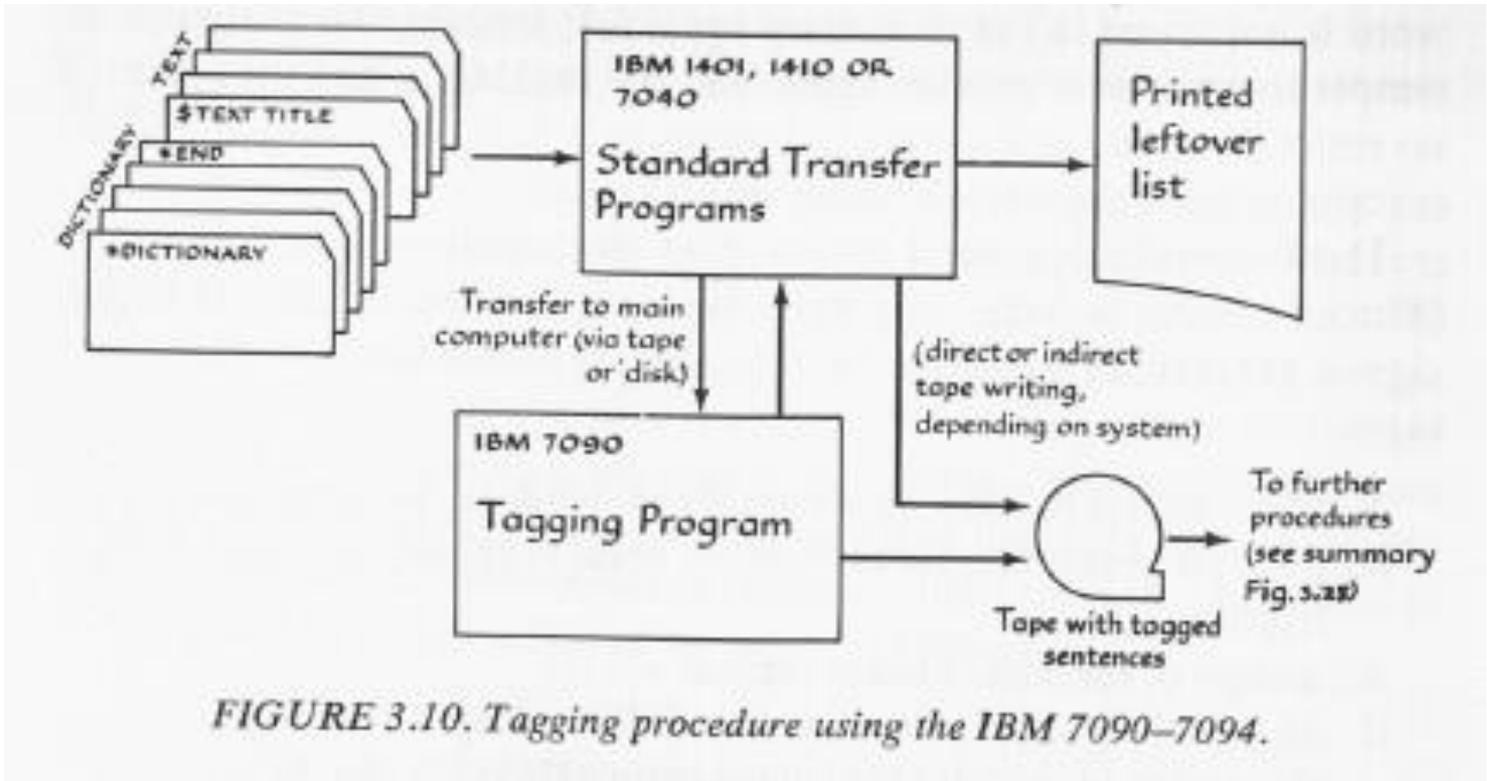
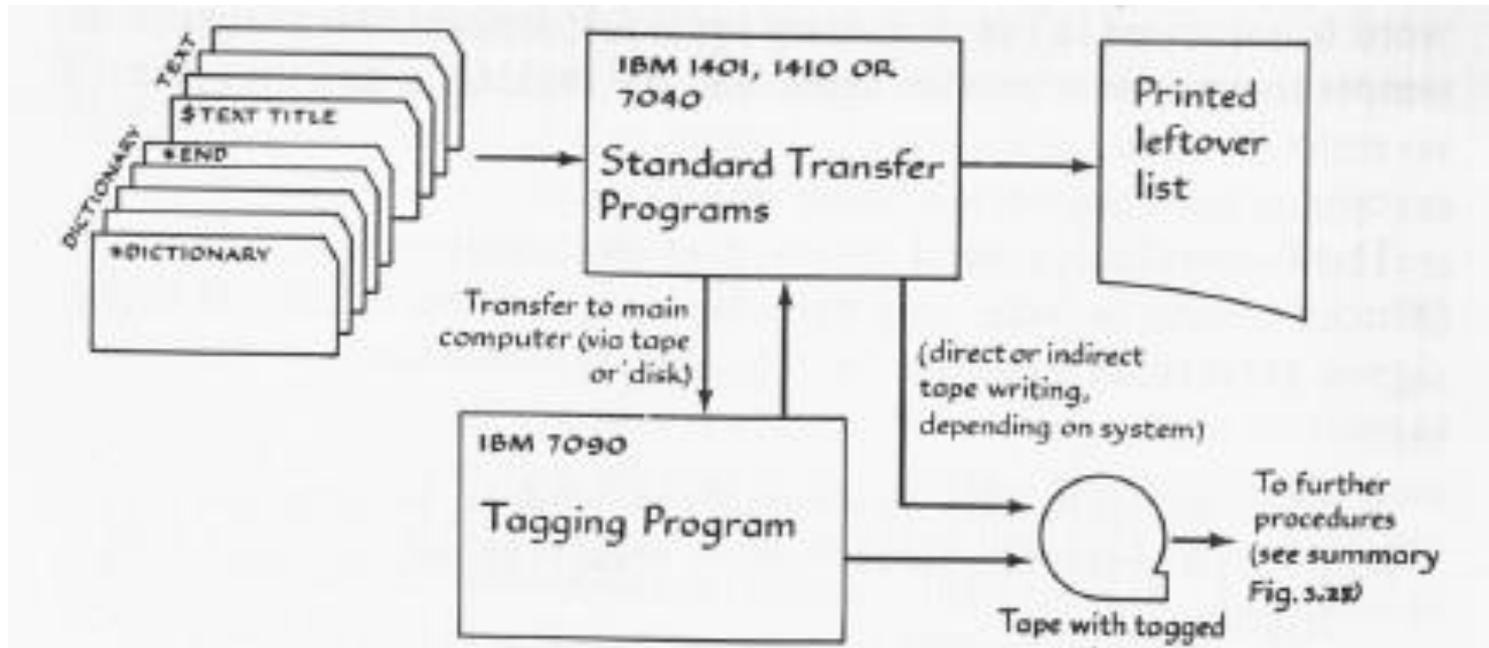


FIGURE 3.10. Tagging procedure using the IBM 7090-7094.

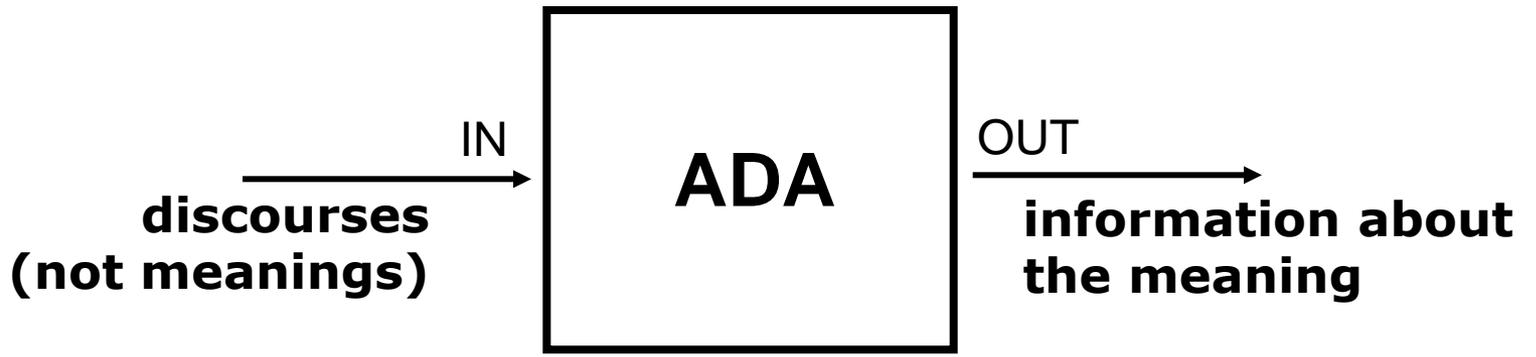


The categories assigned are called *tags*.

A computer scans the text from beginning to end. Each successive text word is looked up in the content analysis dictionary provided by the investigator. The tags assigned by each successfully matched entry word [...] are stored in sequence on a list.

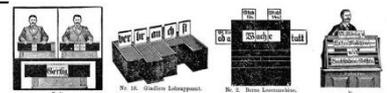
BERNHARD  
DOTZLER

ZU FRÜHEREN  
PROJEKTEN  
DER  
AUTOMATI-  
SIERTEN  
DISKURS-  
ANALYSE



ALGORITHMISCHE

DISKURSANALYSE



the process of production of discourse:

$$\Gamma_x^n \circ L \rightarrow \Delta_x^n$$

In order to reveal  $\Delta_x$  we will attempt to define both the points at which it

the process of production of discourse:

$$\Gamma_x^n \circ L \rightarrow \Delta_x^n$$

In order to reveal  $\Delta_x$  we will attempt to define both the points at which it

**Diskurs:  $D_x$**

**Produktionsbedingungen:  $\Gamma_x$**

**Produktionsprozeß:  $\Delta_x$**

**Äußerungen (utterances):  $U_1, U_2 \dots U_n$**

the process of production of discourse:

$$\Gamma_x^n \circ L \rightarrow \Delta_x^n$$

In order to reveal  $\Delta_x$  we will attempt to define both the points at which it

**Diskurs:  $D_x$**

**Produktionsbedingungen:  $\Gamma_x$**

**Produktionsprozeß:  $\Delta_x$**

**Äußerungen (utterances):  $U_1, U_2 \dots U_n$**

BERNHARD  
DOTZLER

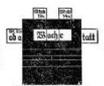
ZU FRÜHEREN  
PROJEKTEN  
DER  
AUTOMATI-  
SIERTEN  
DISKURS-  
ANALYSE

$$U_n = F (DET_1, N_1, V, ADV, P, DET_2, N_2)$$

ALGORITHMISCHE

DISKURSANALYSE

FAKULTÄT MEDIEN, BAUHAUS-UNIVERSITÄT WEIMAR



$$U_n = F (DET_1, N_1, V, ADV, P, DET_2, N_2)$$

	F	DET <sub>1</sub>	N <sub>1</sub>	V	ADV	P	DET <sub>2</sub>	N <sub>2</sub>
U <sub>i</sub> =	a	b	c	d	e	f	g	h
U <sub>j</sub> =	a	b	c	d	j	f	k	m

We now associate a number with the (U<sub>i</sub>, U<sub>j</sub>) pair in each morpho-syntactic category, adopting the convention that two identical terms within the same morpho-syntactic class are represented by the number 1, and different terms by the number 0. Here, for example, we obtain

$$(U_i, U_j) = 1 1 1 1 0 1 0 0$$

$$U_n = F (DET_1, N_1, V, ADV, P, DET_2, N_2)$$

	F	DET <sub>1</sub>	N <sub>1</sub>	V	ADV	P	DET <sub>2</sub>	N <sub>2</sub>
U <sub>i</sub> =	a	b	c	d	e	f	g	h
U <sub>j</sub> =	a	b	c	d	j	f	k	m

We now associate a number with the (U<sub>i</sub>, U<sub>j</sub>) pair in each morpho-syntactic category, adopting the convention that two identical terms within the same morpho-syntactic class are represented by the number 1, and different terms by the number 0. Here, for example, we obtain

$$(U_i, U_j) = 1 1 1 1 0 1 0 0$$

$$U_n = F (DET_1, N_1, V, ADV, P, DET_2, N_2)$$

	F	DET <sub>1</sub>	N <sub>1</sub>	V	ADV	P	DET <sub>2</sub>	N <sub>2</sub>
U <sub>i</sub> =	a	b	c	d	e	f	g	h
U <sub>j</sub> =	a	b	c	d	j	f	k	m

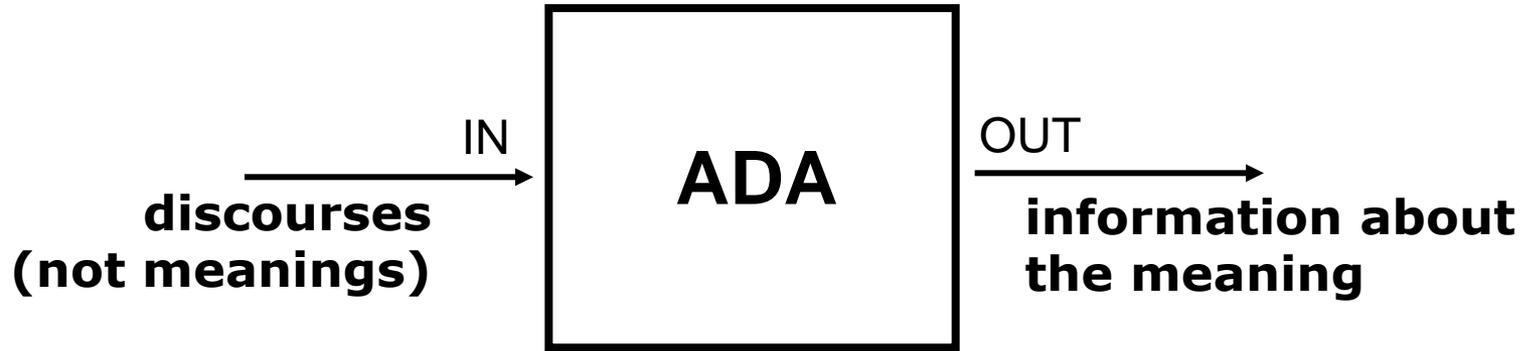
We now associate a number with the (U<sub>i</sub>, U<sub>j</sub>) pair in each morpho-syntactic category, adopting the convention that two identical terms within the same morpho-syntactic class are represented by the number 1, and different terms by the number 0. Here, for example, we obtain

$$(U_i, U_j) = 1 1 1 1 0 1 0 0$$

= Maß der *paradigmatic proximity* zur Bestimmung von *semantic domains*

BERNHARD  
DOTZLER

ZU FRÜHEREN  
PROJEKTEN  
DER  
AUTOMATI-  
SIERTEN  
DISKURS-  
ANALYSE



ALGORITHMISCHE

DISKURSANALYSE

