

Total or Selected Content Analysis¹

Julius Laffal

Connecticut Valley Hospital, Middletown, Conn.

The core of any system of content analysis is a dictionary specifying the content categories to be assigned to verbal units. The system described in this paper utilizes a category dictionary in which all the common words of English are classified into one or two of 116 possible categories.² The associated computer method is at present capable of analyzing a text of 25,000 words in about 20 minutes of machine time, outputting a distribution of frequency of occurrence of the 116 content categories. Up to 15 separate texts may be analyzed at the same time, or the productions of 15 speakers, mixed in any way, as long as each segment is identified with the appropriate source.

The computer analysis program will be discussed briefly, as will some of the considerations which went into the construction of the dictionary. An illustration of the analysis of textual material will be presented, and two studies applying the procedure will be described.

Preparing a text, and machine analysis

For any English language corpus to be analyzed, a simple editing procedure is followed before the text is card punched. The editor

-
1. Presented at the International Conference on Computational Linguistics, Sönga-Säby-Kursgård, Sweden, Sept. 1 - 4, 1969.
 2. Two of the categories (no score and hand-score) are not actually content categories.

eliminates a number of function words (articles, most conjunctions, auxiliary verbs such as to be and to have), changes all verb forms into the present infinitive form, and reduces plural noun forms to the singular. These editing changes are required because the dictionary contains only the root forms of the word. Illustrations 3 and 5 show a text and its edited form. Ultimately, as the dictionary expands to include common forms, many of these editing requirements will be reduced. There are 20,000 root forms in the dictionary at present. The editing procedure reduces an original text by between 15% and 20%: a 25,000 edited text would originally have consisted of around 40,000 words. The computer analysis is directed solely to content, and does not take grammatical form into account. The analysis thus shows what is being talked about, but provides no information that the grammar ordinarily conveys. It may show that aggression is a major theme of a text, but will not show who is aggressing against whom. The researcher may choose to analyze for total content, as represented by the 116 categories, or specify a limited number of content categories, in which case the program will disregard all other categories.

The program is at present written in Fortran IV and MAP languages (ultimately to be entirely in Fortran IV). It consists of a number of subroutines capable of performing the following operations on a text:

1. Print the text itself, numbering each line.
2. Output the text words in order of occurrence, showing speaker and line number.
3. Output the text words in alphabetic order, showing speaker and line number.

4. Output a list of words matched with the stored dictionary, showing the categories which have been applied to each matched word, the speaker number, and the line number.
5. Output a list of words which have more than one possible scoring in the dictionary (ambiguous words), showing either a selected number of the possible scoring, or all the scorings. Speaker number and line number are also indicated.
6. Output a list of words which the dictionary cannot define because they are not represented by dictionary entries.
7. Output profiles of the frequency of occurrence of each of the 116 categories for each speaker.
8. Output a punched deck of category frequency profiles so that further statistical analyses may be applied.

The Dictionary

To find conceptual uniformities in the vocabulary of a language poses formidable problems. In typical content analyses, the researcher takes sentences or larger units, and evaluates them as to the presence or absence of a particular concept, dismissing as irrelevant to his purpose any material which does not bear on the concept. This approach is inadequate if one wishes to deal with all of the content of the sample.

There are certain empirical bases for developing a conceptual system which will deal with all possible contents. The steps in such a process may be described briefly. Numerous words are substantially similar to each other, and essentially synonymous in particular contexts.

Dictionaries take advantage of this fact to introduce us to the meaning of unfamiliar words, by providing a synonymous word which we know. Synonyms are of course not always substitutable for each other without some change of meaning, for words may partake of more than one domain of meaning.

The number of separate domains of meaning each occupied by a few synonyms is extremely large. As a step toward reduction one may group words similar in meaning, even if not synonymous. Words like walk, run, ride, and go readily group together as sharing a common meaning. Application of the criterion of similarity along with that of synonymy leads to broader and more inclusive clusters of words.

A further reduction in the number of clusters or domains of meaning may be effected by applying an additional criterion for grouping words: that of relatedness. If a set of words such as chair, stool, bench, seat, and couch is grouped by synonymy and similarity, considerations of relatedness may also bring table into the same domain of reference as a common item of household furnishing. The merging of table and chair into one group also projects the outlines of a category which could include other words like bed, dresser, bureau, chest, all of which also refer to ordinary household furnishings. These principles of synonymy, similarity, and relatedness, when applied systematically to words, produce a system of categories limited in number but still capable of describing meaningfully all the contents of the language.

When a certain degree of stability has been reached in constructing categories, new words may be considered in terms of how they will fit into the system. This does not preclude changing the system; it

merely weighs all the prior judgments against each new judgment to be made, and accepts change only on compelling grounds. As the system grows, changes are more likely to occur because two existing categories are seen to border too closely on each other and to require redefinition of their boundaries and differences, rather than because new categories are discovered.

A primary consideration in construction of the present conceptual dictionary was that the scheme of language had to provide an adequate basis, at the very least, for discriminating the language of one speaker from another. This meant that each entity in the scheme had to contribute to such potential. In practice this militated against units which were heavily loaded with tokens from everyone's speech and against units which rarely contributed a token to anyone's speech. Such global categories as "nature," "man," "animate" and "inanimate" provide very little basis for discriminating between the word predilections of separate speakers, since they inevitably contribute large numbers of tokens to everyone's speech.

Relatively simple criteria for the acceptability of a category in the overall scheme were (1) that it not draw too many or too few tokens, and (2) that it contribute to discriminating between separate speakers. Where, in the present scheme, a category turned out to be large by these criteria, what was also readily apparent was some basis for splitting the category along one or more lines to produce separate categories. And where a category seemed too narrow, it was possible to merge it with another neighboring category, the domain of reference than swelling to include both meanings. The area of hostility and aggression was

one which required subdivision along lines which differentiated violence, anger, and simple disagreement. The area of sex, in which at the outset a discrimination was provided between heterosexual and homosexual, ended by becoming a single area, since very few references occur to homosexuality as such as opposed to sexuality in general.

One other criterion which was applied is related to well known ideas of reliability and validity. The domains of reference in a scheme of language must be sufficiently discriminable from each other so that separate judges considering a word will be in agreement as to which categories are involved. If the domains of meaning are too closely related, this will be a difficult judgment and an unreliable one. To the extent that it is unreliable, its validity and its pertinence to other behavior, will be dubious. Distinctiveness of a domain of reference in a dictionary scheme is therefore an important desideratum.

Over the years, as texts from various sources were analyzed, with a continuing effort to fit new words into existing groups or to redefine the word groups to make them more congruent with word meanings, a dictionary of words with their appropriate categories emerged. The dictionary and its mode of application are illustrated on the following pages which show:

1. Definitions of a few categories.¹
2. A page of dictionary entries showing associated categories, and illustrative sentences.

1. A detailed description of the categories is contained in Laffal, J. Pathological and normal language. New York: Atherton Press, 1965.

3. A selection from Jonathan Swift's Gulliver's Travels.
4. Handscoring of the Gulliver's Travels selection.
5. Selection from Gulliver's Travels, edited and punched on cards for machine analysis.
6. Profiles derived by computer and by hand analysis of the Gulliver's Travels selection. These profiles also show an alphabetic listing of the categories by heading word. In the computer analysis the frequency profile is updated by the different categories in the first two meanings for each ambiguous word, whereas in the hand analysis, only the categories in the relevant meaning are applied.

1. Definitions of a Few Categories

CHANGE: The kinds of words which are scored here are:

- a. Words which refer to deviation: bent, crooked, jagged, tortuous, winding, derail, avoid, evade, distract. "He turned (CHANGE, GO) the corner."
- b. Words referring to alteration either of form or position: convert, mutation, revise, transform, shift. "The dress was altered (CHANGE)."
- c. Words referring to instability or to inconstancy. "His opinion fluctuates (CHANGE) from moment to moment."
- d. Words referring to transience, temporariness, or momentariness. "I caught a fleeting (CHANGE, FAST) glimpse of him as he ran." "The country was ruled by a provisional (CHANGE) government."

CLEAN: The types of words scored here are:

- a. Words referring to personal cleanliness and grooming: toothbrush, comb, wash, towel, soap. "Scrub (CLEAN) behind your ears."
- b. Words which refer to elimination of dirt: purification, cleanliness, dusting, laundry. "She finished mopping (CLEAN) the floor."

CLOTHING: All references to wearing apparel and things pertinent to being clothed or dressed are scored here. In addition, related activities or references, such as sew and darn, are scored here. Some clothing is clearly male, female, or infant. In such cases the additional scoring of MALE, FEMALE, or YOUNG is applied. "His trousers (CLOTHING, MALE) had to be altered." "She wore (CLOTHING) a beautiful dress (CLOTHING, FEMALE)." "She changed the baby's diaper (CLOTHING, YOUNG)."

COLD: All words in which the idea of coolness or coldness is of importance are scored here. "It's freezing (COLD) in here."

2. A Page of Dictionary Entries Showing Associated Categories
and Illustrative Sentences*

COED	N FEMALE	REASON 2	An attractive <u>coed</u> won the contest.
COEDUCATION	N JOIN	REASON 2	<u>Coeducation</u> is now common in college.
COEFFICIENT	N NUMBER		A <u>coefficient</u> is a multiplier.
COEQUAL	J AGREE 3		The soldiers were <u>coequal</u> in rank.
COERCE	V CONFL. 1	DOM. 1	Bandits <u>coerce</u> the travelers.
COEVAL	J AGREE 3	TIME 4	Roosevelt and Churchill were <u>coeval</u> leaders.
COEXIST	V JOIN	LIVING	Nations must <u>coexist</u> peacefully.
COFFEE	N DRINK	VEGETA.	He drank <u>coffee</u> in the morning.
COFFEE	N COLOR		<u>Coffee</u> is a brown color.
COFFER	N HOLLOW 2		She kept her jewels in a <u>coffer</u> .
COFFIN	N DEAD	HOLLOW 2	A flag was draped on the <u>coffin</u> .
COG	N HILL		The <u>cog</u> of one gear fits the other.
COG	N FALSE		His unfulfilled promise was a <u>cog</u> .
COG	V HILL	JOIN	The carpenter will <u>cog</u> the boards.
COGENCY	N REASON 1	TRUE	The argument has great <u>cogency</u> .
COGENT	J REASON 1	TRUE	The debater presented a <u>cogent</u> argument.
COGITATE	V REASON 1		It is wise to <u>cogitate</u> before acting
COGNAC	N DRINK	VEGETA.	<u>Cognac</u> is a French brandy.
COGNATE	J AGREE 3	JOIN	Father and vater are <u>cognate</u> words.
COGNITION	N REASON 1		Learning is the basis of <u>cognition</u> .
COGNIZANCE	N REASON 1		The court took <u>cognizance</u> of the dispute.

*J, adjective; N, noun; V, verb.

3. A Selection from Jonathan Swift's Gulliver's Travels

My father had a small estate in Nottinghamshire; I was the third of five sons. He sent me to Emmanuel College in Cambridge at fourteen years old, where I resided three years, and applied myself close to my studies; but the charge of maintaining me (although I had a very scanty allowance) being too great for a narrow fortune, I was bound apprentice to Mr. James Bates, an eminent surgeon in London, with whom I continued four years, and my father now and then sending me small sums of money, I laid them out in learning navigation and other parts of the mathematics useful to those who intend to travel, as I always believed it would be some time or other my fortune to do. When I left Mr. Bates, I went down to my father, where, by the assistance of him and my uncle John and some other relations, I got forty pounds, and a promise of thirty pounds a year to maintain me at Leyden; there I studied physic two years and seven months, knowing it would be useful in long voyages.

Soon after my return from Leyden, I was recommended by my good master Mr. Bates to be surgeon to the Swallow, Captain Abraham Pannell Commander, with whom I continued three years and a half, making a voyage or two into the Levant and some other parts. When I came back I resolved to settle in London, to which Mr. Bates, my master, encouraged me, and by him I was recommended to several patients. I took part of a small house in the Old Jewry, and, being advised to alter my condition, I married Mrs. Mary Burton, second daughter to Mr. Edmund Burton, hosier in Newgate Street, with whom I received four hundred pounds for a portion.

But my good master Bates dying in two years after, and I having few friends, my business began to fail; for my conscience would not

suffer me to imitate the bad practice of too many among my brethren. Having, therefore, consulted with my wife and some of my acquaintances, I determined to go again to sea. I was surgeon successively in two ships, and made several voyages for six years to the East and West Indies, by which I got some addition to my fortune. My hours of leisure I spent in reading the best authors, ancient and modern, being always provided with a good number of books, and when I was ashore in observing the manners and dispositions of the people, as well as learning their language, wherein I had a great facility by the strength of my memory.

The last of these voyages not proving very fortunate, I grew weary of the sea, and intended to stay at home with my wife and family. I removed from the Old Jewry to Fetter Lane, and from thence to Wapping, hoping to get business among the sailors, but it would not turn to account. After three years' expectation that things would mend, I accepted an advantageous offer from Captain William Prichard, master of the Antelope, who was making a voyage to the South Sea. We set sail from Bristol, May 4, 1699, and our voyage at first was very prosperous.

OUT

LITTLE ALL 1 MONEY SR MONEY SEPARATE REASON 2 GO
 small sums of money, I laid them out in learning navigation,
 CONF.3 SOME NUMBER HELP WANT 1
 and other parts of the mathematics, useful to those who intend to
 AGREE 2
 GO AGREE 3 SR DURABLE REASON 1 TIME 4
 travel; as I always believed it would be some time or
 END GO CURE
 CONF.3 SR HELP TIME 4 SR SEPARATE MALE SHARP 1 SR
 other my fortune to do. When I left Mr. Bates, I
 DOM.2 DOM.2
 GO DOWN SR MALE PLACE HELP MALE SR
 went down to my father; where, by the assistance of him and my
 JOIN
 MALE SOME CONF.3 JOIN SR POSSESS NUMBER MONEY
 uncle John, and some other relations, I got forty pounds,
 DURABLE
 AGREE 2 NUMBER MONEY TIME 4 HELP SR NEAR PLACE
 and a promise of thirty pounds a year to maintain me at Leyden:
 SR REASON 1 REASON 3 NUMBER TIME 4 NUMBER TIME 4
 there I studied physic two years and seven months,
 BIG
 REASON 1 REASON 3 HELP DURABLE GO
 knowing it would be useful in long voyages.

NEAR BACK AGREE 2
 TIME 3 BACK SR GO SEPARATE PLACE SR GOOD
 Soon after my return from Leyden, I was recommended,
 CURE CURE ANIMAL SEA
 SR GOOD DOM.2 MALE SHARP 1 SHARP 1 UP TRANS.
 by my good master Mr. Bates, to be surgeon to the Swallow (ship),
 DOM.2 DOM.2 JOIN SR DURABLE NUMBER
 Captain Abraham Pannell commander; with whom I continued three
 NUMBER
 TIME 4 SOME GO GO NUMBER IN PLACE
 years and a half, making a voyage or two into the Levant, and
 DURABLE
 SOME CONF.3 SOME TIME 4 SR GO BACK SR WANT 1 LIVING
 some other parts, When I came back, I resolved to settle
 CURE
 IN PLACE MALE SHARP 1 SR DOM.2 HELP SR
 in London, to which Mr. Bates, my master, encouraged me, and by
 CURE AGREE 2
 SHARP 1 SR GOOD SOME ILLNESS SR POSSESS SOME
 him I was recommended to several patients. I took part of
 GROUP HELP
 LITTLE HOLLOW 1 IN TIME 1 LAW PLACE REASON 1
 a small house in the Old Jury (Place); and being advised
 JOIN FEMALE
 CHANGE SR HAPPENING SR SEX FEMALE JOIN NUMBER
 to alter my condition, I married Mrs. Mary Burton, second

FEMALE CLOTHING CLOTHING
 JOIN? MALE COMMERCE COMMERCE IN GO PLACE
 daughter to Mr. Edmund Burton, hosier in Newgate-street (Place);
 JOIN SR POSSESS NUMBER MONEY AGREE 3 SOME
 with whom I received four hundred pounds for a portion

CURE

CONF.3 SR GOOD DOM.2 SHARP 1 DEAD NUMBER TIME h
 But, my good master Bates dying in two years
 BACK SR POSSESS SOME AGREE 1 SR COMMERCE BEGIN SUB.1
 after, and I having few friends, my business began to fail;
 GOOD AGREE 2 AGREE 3
 SR REASON 1 NEGATION UPSET SR FALSE BAD
 for my conscience would not suffer me to imitate the bad

JOIN

DURABLE SHARP 2 ALL 2 JOIN SR MALE
 practice of too many among my brethren. Having therefore
 HELP FEMALE DOM.1
 REASON 1 JOIN SR SEX SOME SR NEAR SR REASON 1
 consulted with my wife, and some of my acquaintance, I determined

CURE

SEA
 GO AGREE 3 SEA SR SHARP 1 FORWARD IN NUMBER TRANS.
 to go again to sea. I was surgeon successively in two ships,
 GO SOME GO NUMBER TIME h PLACE PLACE
 and made several voyages, for six years, to the East and West.

22

OMIHTOIO BIG OMIHTOIO ALL 2 ELIHTY
 PLACE OO SR POSSESSO SOME O JOIN BOPTMSR MONEY SHAGU
 Indies; by which I got some addition to my fortune. My
 CALM UNION LANGUAGE 2 DOM.109
 TIME h SLOW SR WANT 2 SEE GOOD LANGUAGE 2
 hours of leisure I spent in reading the best authors,
 TIME 2 DURABLE HELP GOOD NUMBER
 ancient and modern; being always provided with a good number of
 EARTH DURABLE
 LANGUAGE 2 TIME h SR SEA SEE HAPPENING
 books; and when I was ashore, in observing the manners and
 WANT 1 GROUP AGREE 3 REASON 2 GROUP LANGUAGE 1
 dispositions of the people, as well as learning their language,
 DOM.1
 SR POSSESS BIG EASY DOM.1 SR REASON 1
 wherein I had a great facility by the strength of my memory.
 END GO NEGATION TRUE SHARP 2
 The last of these voyages not proving very
 SUB.1
 HELP SR CHANGE UPSET SEA WANT 1 DURABLE NEAR
 fortunate, I grew weary of the sea, and intended to stay at
 FEMALE GROUP
 HOLLOW 1 JOIN SR SEX SEPT JOIN SR SEPARATE SEPARATE
 home with my wife and family. I removed from the

GROUP

TIME 1 LAW PLACE CONFINE GO PLACE SEPARATE
 Old Jury (Place) to Fetter - Lane (Place), and from thence
 PLACE WANT 1 POSSESS COMMERCE JOIN SEA CONF.3
 to Wapping, hoping to get business among the sailors; but it
 CHANGE
 NEGATION GO HELP BACK NUMBER TIME 4 WANT 1
 would not turn to account. After three years expectation
 AGREE 2
 HELP SR POSSESS HELP HELP
 that things would mend, I accepted an advantageous offer from
 SEA
 DOM.2 DOM.2 ANIMAL TRANS. GO
 Captain William Prichard, master of the Antelope (ship), who was making
 GO
 a voyage to the South - Sea. We set sail from Bristol May
 NUMBER FORWARD DOM.1
 NUMBER TIME 4 SR GO NUMBER SHARP 2 HELP
 4, 1699, and our voyage at first was very prosperous.

5. Selection from Gulliver's Travels

Edited and Punched on Cards for Machine Analysis

MY FATHER HAVE SMALL ESTATE IN NOTTINGHAMSHIRE PLACE, I(P)* WAS THE THIRD OF FIVE SON. FATHER SEND ME TO EMANUEL COLLEGE IN CAMBRIDGE PLACE FOURTEEN YEAR OLD, WHERE I(P) RESIDE THREE YEAR, AND APPLY MYSELF CLOSE TO MY STUDY, BUT THE CHARGE OF MAINTAIN ME ALTHOUGH I(P) HAVE VERY SCANT ALLOWANCE BEING(V) TOO GREAT FOR NARROW FORTUNE, I(P) WAS BOUND APPRENTICE TO MR. JAMES-BATES SURGEON, AN EMINENT SURGEON IN LONDON PLACE, WITH WHOM I(P) CONTINUE FOUR YEAR, AND MY FATHER NOW AND THEN SEND ME SMALL SUM OF MONEY, I(P) MONEY OUT LEARN NAVIGATE, AND OTHER PART OF THE MATHEMATICS, USEFUL TO THOSE WHO INTEND TO TRAVEL, AS I(P) ALWAYS BELIEVE IT WOULD BE SOME TIME OR OTHER MY FORTUNE TO DO(V) WHEN I(P) LEAVE MR. BATES SURGEON, I(P) GO DOWN TO MY FATHER, WHERE, BY THE ASSIST OF FATHER AND MY UNCLE JOHN, AND SOME OTHER RELATE, I(P) GET FORTY POUND, AND PROMISE OF THIRTY POUND YEAR TO MAINTAIN ME AT LEYDEN PLACE, THERE I(P) STUDY PHYSICS TWO YEAR AND SEVEN MONTH, KNOW PHYSICS WOULD BE USEFUL LONG VOYAGE.

SOON AFTER MY RETURN FROM LEYDEN PLACE, I(P) WAS RECOMMEND, BY MY GOOD MASTER MR. BATES SURGEON, TO BE SURGEON TO THE SWALLOW SHIP, CAPTAIN ABRAHAM-PANNELL COMMAND, WITH WHOM I(P) CONTINUE THREE YEAR AND HALF, MAKE VOYAGE OR TWO INTO THE LEVANT PLACE, AND SOME OTHER PART.

*Letters in parentheses indicate part of speech. Where word may have more than one part of speech, specification of the part selects the appropriate dictionary entry.

WHEN I(P) COME BACK, I(P) RESOLVE TO SETTLE IN LONDON PLACE, TO WHICH MR. BATES SURGEON, MY MASTER, ENCOURAGE ME, AND BY SURGEON I(P) WAS RECOMMEND TO SEVERAL PATIENT. I(P) TAKE PART OF SMALL HOUSE IN THE OLD JURY PLACE, AND BEING(V) ADVISE TO ALTER MY CONDITION, I(P) MARRY MRS. MARY-BURTON DAUGHTER, SECOND DAUGHTER TO MR. EDMUND-BURTON HOSIER, HOSIER IN NEWGATE STREET PLACE, WITH WHOM I(P) RECEIVE FOUR POUND FOR PORTION.

BUT, MY GOOD MASTER BATES SURGEON DIE(V) TWO YEAR AFTER, AND I(P) HAVE FEW FRIEND, MY BUSINESS BEGIN TO FAIL, FOR(C) MY CONSCIENCE WOULD NOT SUFFER ME TO IMITATE THE BAD PRACTICE OF TOO MANY AMONG MY BRETHREN. THEREFORE CONSULT WITH MY WIFE, AND SOME OF MY ACQUAINTANCE, I(P) DETERMINE TO GO AGAIN TO SEA. I(P) WAS SURGEON SUCCESSIVE IN TWO SHIP, AND MAKE SEVERAL VOYAGE, FOR SIX YEAR, TO THE EAST AND WEST INDIES PLACE, BY WHICH I(P) GET SOME ADDITION TO MY FORTUNE. MY HOUR OF LEISURE I(P) SPEND READ THE BEST AUTHOR, ANCIENT AND MODERN, BEING(V) ALWAYS PROVIDE WITH GOOD NUMBER OF BOOK, AND WHEN I(P) WAS ASHORE, OBSERVE THE MANNER AND DISPOSE OF THE PEOPLE, ASWELLAS LEARN PEOPLE LANGUAGE, WHEREIN I(P) HAVE GREAT FACILITY BY THE STRENGTH OF MY MEMORY.

THE LAST OF THESE VOYAGE NOT PROVE VERY FORTUNATE, I(P) GROW WEARY OF THE SEA, AND INTEND TO STAY AT HOME WITH MY WIFE AND FAMILY. I(P) REMOVE FROM THE OLD JURY PLACE TO FETTER LANE PLACE, AND FROM THENCE TO WAPPING PLACE, HOPE TO GET BUSINESS AMONG THE SAILOR, BUT IT WOULD NOT TURN TO ACCOUNT. AFTER THREE YEAR EXPECT THAT THING WOULD MEND, I(P) ACCEPT AN ADVANTAGE OFFER FROM CAPTAIN WILLIAM-PRICHARD, MASTER OF THE ANTELOPE SHIP, WHO WAS MAKE VOYAGE TO THE SOUTH SEA. WE SET SAIL

FROM BRISTOL PLACE MAY FOUR, SIX, AND OUR VOYAGE AT FIRST WAS VERY
PROSPER.

6. Profiles Derived by Computer and by Hand Analysis
of the Gulliver's Travels Selection

(S1 is the machine output profile, S2 the hand scoring profile).

	S1	S2		S1	S2		S1	S2
ABSURD	0	0	COMMERCE	5	4	FAST	0	0
AGREE 1	1	1	CONFIN	6	1	FEATURE	0	0
AGREE 2	14	6	CONFLICT 1	0	0	FEMALE	5	5
AGREE 3	15	6	CONFLICT 2	1	0	FORWARD	2	2
ALL 1	1	1	CONFLICT 3	10	8	FUNCTION	1	0
ALL 2	2	3	CRIME	0	0	GO	25	22
ALL 3	0	0	CURE	9	9	GOOD	9	7
ANIMAL	2	2	DEAD	1	1	GROUP	7	5
ART	0	0	DIFFICULT	0	0	HAND SCORE	2	*
ASTRONOMY1	0	0	DIRTY	0	0	HAPPENING	2	2
ASTRONOMY2	0	0	DOMINANCE1	15	6	HEAR	0	0
BACK	6	5	DOMINANCE2	12	12	HELP	19	17
BAD	1	1	DOWN	1	1	HILL	0	0
BEGIN	1	2	DRINK	0	0	HOLLOW 1	9	4
BIG	5	4	DURABLE	11	11	HOLLOW 2	0	0
BLURRED	3	0	EARTH	2	2	HOT	0	0
CALM	1	1	EASY	1	1	HOUSEHOLD	0	0
CHANGE	3	3	EAT	1	0	ILLNESS	1	1
CLEAN	0	0	END	5	2	IN	8	8
CLOTHING	3	2	ESSENTIAL	1	0	INDIVIDUAL	4	0
COLD	0	0	FALSE	1	1	JOIN	18	16
COLOR	0	0	FAR	0	0	LANGUAGE 1	2	1

LANGUAGE 2	3	3	POSSESS	13	11	TIME 1	5	4
LAW	3	2	REASON 1	14	9	TIME 2	2	2
LITTLE	5	5	REASON 2	4	4	TIME 3	1	1
LIVING	2	2	REASON 3	2	2	TIME 4	21	17
MALE	13	13	SACRED	5	0	TRANSPORTA	3	3
MATERIAL	1	0	SEA	9	9	TRIVIAL 1	0	0
MEASUREMEN	3	0	SEE	2	2	TRIVIAL 2	0	0
MECHANISM	0	0	SELF REFER	58	58	TRUE	1	1
MONEY	6	9	SEPARATE	9	7	UNREAL	0	0
NEAR	12	5	SEX	3	3	UP	1	1
NEGATION	3	3	SHARP 1	10	9	UPSET	4	2
NO SCORE	162	*	SHARP 2	10	5	VEGETATION	0	0
NUMBER	24	23	SLOW	1	1	WANT 1	13	8
OPEN	0	0	SOME	13	12	WANT 2	2	1
OUT	1	1	STRUCTURE	5	0	WORK	1	0
PLACE	19	19	SUBMISS 1	3	3	YOUNG	0	0
PLAY	2	0	SUBMISS 2	0	0			
PROFILE TOTALS						731 446		

*Hand Score and No Score items are not tabulated when a profile is scored by hand.

To illustrate the application of the computer analysis, two studies are described, one, of the free flowing language of children, the other, of scientific writings.

A book by Evelyn G. Pitcher and Ernst Prelinger (1963) presents stories told by boys and girls from ages 2 to 5, taken down in shorthand by one of the authors. These were a group of economically privileged children attending a nursery and kindergarten of a private school. Most of the children were of superior or high average intelligence. When the occasion presented itself and the child was for the moment playing alone, the researcher asked, "Tell me a story." Only those stories were rejected which were a retelling of a familiar fairy tale or television show.

The stories taken from the Pitcher-Prelinger book were transcribed onto IBM cards and analyzed by computer to obtain profiles showing the distribution of the content of each age group. In order to sharpen the subsequent analysis, all categories showing a frequency of less than 1% of the total responses within each profile were eliminated. Also eliminated was the one category (ANIMAL) which showed a uniform high frequency (over 7% of the total) for each profile. These extremely low and extremely high frequency categories are eliminated because they contribute only to raising all correlations between profiles and do not contribute to discriminating between profiles. The profiles, now consisting of 52 categories were then correlated with each other, and the matrix of correlations factor analyzed. I will not present the actual factor analysis, but only describe the outcome.

The first factor in the analysis is a "female" factor. The girls in all age groups load very highly on it. When we look in detail at the categories which contribute most heavily to this factor we find that YOUNG, FEMALE, and HELP are the three outstanding contributors. In the category system, FEMALE and HELP are the two scorings which apply to the word mother and its variants. These findings suggest that the girls of the study had as central themes in their fantasy, ideas relating to mothers (feminine helpers) and children.

The second factor is more clearly relevant to boys, with ages 3 and 4 showing predominance. The most significant categories here are GO, MALE, and TRANSPORTATION. Thus, the boys in the group, particularly those in the 3 and 4 year bracket, had prominently in their fantasies, references to maleness, physical movement, and vehicles of transportation.

The third factor applied primarily to girls age 4 and boys age 5. Primary categories were MALE, GO, and FEATURE (bodily features). The FEATURE and MALE categories would be consistent with a psychoanalytic view that at this age both boys and girls are in the oedipal period and are becoming preoccupied with sexuality, and particularly with the relationship to father.

The fourth factor seems most clearly applicable to boys age 2, although to a lesser degree it is characteristic of boys age 3, suggesting a decay process with age. Here the outstanding categories are CONFLICT 1 (physical harm), GO, and YOUNG. Like the other boys, the two year olds were also occupied with physical movement. Like the

girls, they were occupied with ideas about children; but violent aggression was their most prominent concern and distinguishes them most clearly from other groups.

A second study by Hartsough and Laffal explores the question whether scientists who have been described as visual imagists and scientists who have been described as verbal imagists (Roe, 1952) will also differ in the types of content employed in their writing. The four sciences which Roe had studied were Social Science, Biology, Experimental Physics, and Theoretical Physics. Based on a series of projective tests and interviews with scientists in these fields, as well as their self-reports of problem solving activities and typical modes of thinking about problems, she concluded that Social Scientists and Theoretical Physicists were verbal imagists, while Biologists and Experimental Physicists were visual imagists. Verbal imagery is thinking in words or talking to oneself, while visual imagery is thinking in pictures.

In our study, we selected three outstanding living scientists from each of these four disciplines, the Social Sciences being represented by Anthropologists. For each scientist one representative book was chosen from his technical publications addressed to colleagues, and one from his popular writings for the laity. The key words science, scientist, Biology, Anthropology, and Physics were used to identify contexts for comparison. A context included the sentence containing the key word, the preceding sentence, and the subsequent sentence. In

addition to the books of the scientists, an Anthropology, a Biology and a Physics textbook of this time were also selected. The content of the textbooks was sampled by taking three sentences alternately from the top, middle and bottom of every other page, without regard to any key words. The writings of the various scientists, and in the texts, were analyzed by the computer program described, and content category profiles were obtained. These profiles were then correlated with each other and factor analyzed. I will not present details of the factor analysis, but only the results.

Highly represented on the first factor were all of the Anthropological writings, and the writings of the Theoretical Physicists. Since these two groups together constitute what Roe had called the verbal scientists, the first factor could be considered a verbal factor. The actual content which was stressed in the writings of these scientists, however, evokes more the idea of a "humanistic" factor, since the particular categories contributing most importantly to the factor are REASON 1 (knowledge and thinking), GROUP, HELP, and GOOD.

Thus, the content analysis tells us that what manifests itself as a highly developed verbal communicative skill, also reflects a concern with human affairs and human issues. Since language is par excellence the social skill, it is natural that those who use it best would also be those whose attention is directed toward the human condition.

The second factor was clearly related to Physics, the writings of the Theoretical and Applied Physicists loading highest on it. The categories which contributed most to this factor were: REASON 3

(science), MATERIAL, REASON 1 (knowledge and thinking) and SOME. Other categories which were stressed were ESSENTIAL and TIME 4 (general time). The third factor was clearly a biological factor, since the biological writings loaded highest on it. The categories of greatest importance were: FUNCTION (bodily functions), BEGIN CHANGE, MATERIAL and FORWARD.

To be noted is that no "visual" factor as such appeared in the analysis. This suggests that the contrast which had been drawn by Roe between "verbal" and "visual" may actually be a contrast between those with a central humanistic concern inclined to be strongly verbal in their thinking, and those who do not have this same central interest, who are more likely to stress non-verbal processes in their thinking.

The method of total content analysis which I have described is applicable to many language content problems. Beside the applications described in the present paper, the technique has been used in studies of the symbolism of key words in the language of a psychotic patient (Laffal, 1960); changes in the language of a patient in treatment (Laffal, 1961); comparison of the language of therapists in talking about different patients (Watson & Laffal, 1963); comparisons of the content given by subjects in single word association, continuous word association, and free speech in response to the same stimulus word (Laffal & Feldman, 1962, 1963); comparison of the language of two and three speakers in conversations about a variety of topics (Laffal, 1967).

In broadest terms the method may be applied to comparison of one individual with himself (for example, over time); one individual with

another; a group with itself or with other groups; the contexts surrounding one topic with those surrounding some other topic. The unique contribution which it makes is in the use of a category dictionary which classifies all the common words of English into one or two content categories, so permitting rapid display either of all the content or of selected contents, in a text.

References

- Laffal, J. The contextual associates of sun and God in Schreber's autobiography. J. abnorm. soc. Psychol., 1960, 61, 474-479.
- Laffal, J. Changes in the language of a schizophrenic patient during psychotherapy. J. abnorm. soc. Psychol., 1961, 63, 422-427.
- Laffal, J., & Feldman, S. The structure of single word and continuous word associations. J. verb. Learn. and verb. Behav., 1962, 1, 54-61.
- Laffal, J., & Feldman, S. The structure of free speech. J. verb. Learn. and verb. Behav., 1963, 2, 498-503.
- Laffal, J. Characteristics of the three person conversation. J. verb. Learn. and verb. Behav., 1967, 6, 555-559.
- Roe, Anne. The making of a scientist. New York: Dodd, Mead, 1952.
- Watson, D. L., & Laffal, J. Sources of verbalizations of psychotherapists about patients. J. gen. Psychol., 1963, 68, 89-98.