

The Headline-Project

A Cross-Cultural Approach for the Development of
Textual Indicators

Alexander Deichsel
University of Hamburg
Seminar für Sozialwissenschaften

Preliminary Paper for the
Conference on Cross-National Research
in the Social Sciences
Institute for Social Research
The University of Michigan, Ann Arbor, Oct. 2-8, 1977

THE HEADLINE-PROJECT

A Cross-cultural Approach for the Development of Textual Indicators

I. RESEARCH ORIENTATIONS

In the course of discussion at the Pisa meeting the plan emerged to use content analytical research technology for dealing with questions raised by social indicator research groups. The results and problems discussed at Pisa demonstrated the possibilities of computer aided content analysis to generate primary data from mass media material in an extensive manner. In spite of all restrictions it was accepted to start a textual indicator project. To begin with, the interested participants agreed in concentrating on newspaper headlines because of their simple syntactical form, their validity as regards public attention structure, and their easy accessibility.

But what are the characteristics of a social indicator project compared to 'normal' content analytical investigations? The differences are gradual not essential ones:

- it demands continuous inquiries
- it is a long term analysis with a 'historical' time dimension
- it is less orientated to test theories than to describe social or political opinion-climates by more or less theoretically supported social reporting
- its research orientations are connected with general indicator problems, as regards the normative and political aspects as well as the theoretical and data production ones.

The result of these considerations at Pisa has been a first draft of the project in 1975 (1).

In the meantime the project has been coordinated at Hamburg University. Some difficulties had to be recognized and put into account : Because of the amount of textual material in time series it was decided to use automatic coding procedures. Though computer programs and dictionaries are dramatically different between interested countries, the headlines should not be translated (i.e. into English, because of the sophisticated instruments available for this language), but analysed in their original language. The project should be organized around modest research questions, so that researchers without dictionaries and programs at their disposal might be able to gather home-made beginnings around the project.

How to start the project ? After some hesitations the Hamburg group decided to collect a certain amount of European headlines and to compute them at Hamburg rather than to wait for a decentralized organizational network. Simultaneously contacts to interested persons were developed and the Hamburg research material may be outthanded to national researchers if wanted. As to the problem of what indicator(s) should be set up, we decided pragmatically to begin with the following question : What are the cross-cultural interrelations observed in newspaper headlines of different countries and intra-national regions ? Thus we expected to give a life to the project without rejecting any plan of any future co-researcher.

II. PROBLEMS UNDER DISCUSSION

Several problem areas are being discussed by correspondences, informal meetings and teaching courses.

Social Theories Concerning Headlines

A survey of nearly all available literature concerning headline analysis was done by a student group at Annenberg School of Communication under the guidance of K. KRIPPENDORFF. A general approach formulating definitions and isolating criteria to decide whether a given text is a headline or not is on its way at Annenberg School as well. Studies using automatic coding on German headlines and television news are in progress at Hamburg University. Although some colleagues of the Hamburg research group wanted to extend the text area in order to compare headlines to their articles and though others pointed at the phenomenon of 'headline-blindness' among certain people and under certain circumstances, all of us shared the premise : that headlines bear characteristics allowing to describe certain elements of public orientations and aspirations, that they permit to detect certain socio-cultural climates, thematic structures and inter-cultural patterns. Those data may then be interpreted in different models concerning public opinion as such, its causes and consequences. Thus headlines have been accepted as useful observation units for social reporting purposes.

Sampling Cultures, Newspapers and Headlines

For the time being the sampling^{of} countries (i.e. cultures and languages) is determined brutally by the combination of the project group. Although it may be possible to observe headlines of 'foreign' countries - and this is

what the Hamburg group does at the moment -, the aim is to cooperate with groups who work on their own culture with a cross-cultural research orientation. The choice of newspapers in a specific country should be decided by the interested researchers (special interests; organizational structure of mass-media; variety of newspaper types). As to the headlines it was suggested to start with the front page headlines. The decision whether to examine only the most important one or all of them should be taken by the group in each country (other arrangements of the front page may demand adequate solutions). At Hamburg we began with the most important headline of the front page (singled out according to special criteria) (2).

Categories

Although the general discussion on 'relevant' content dimensions for text indicators has not yet taken place, there has been an orientation towards four basic variables : ACTORS, PROCESSES, THEMATIC AREAS, VALUES. We presume those aspects to be of general interest. They may constitute a nuclear conceptual realm. They may be completed by other categories in connection with special research questions as well as they may be splitted and diversified by sub-categories. Most of the social science projects on media content include those four dimensions. We expect that the data generated by this instrument will constitute a platform for first steps and further developments. Last not least for manual content research those variables may be coded easily even.

Dictionaries

Although manual content analysis should not be excluded a priori, in the long run the aims of the project call for automatic coding. This technique demands dictionaries. Some participants must begin with writing one for the project thus preparing an instrument which can be used and completed in further studies about media content in their languages. For these researchers the first step will be a word-by-word dictionary as used at present for German headlines. A later development will be the introduction of disambiguation as it works in the American HARVARD and VALUE dictionaries.

An examination was done at Hamburg of all available dictionaries and more private tag lists. All of them contain the four basic variables mentioned above (although they are not identic). A first version of a 'Headline-Dictionary' will be operative in spring '78 and can be mailed if ordered.

Computer Programs

We suggest to decentralize the computer work too at a later date. This means that the analysis of English and American headlines can be done in more detail than that of the German ones or of other languages. However, a minimum of consensus is expected so that the data may be comparable. This strategy will permit researchers to adapt the systems at their disposal to content analytical social research purposes. A guide to computer systems for content analysis on an international level is on its way (3). The computing of headlines in Hamburg is done by the EVA-system (without disambiguation). EVA is written in ALGOL 60/TAS.

III. THE PRESENT 'STATE OF AFFAIRES'

Four question areas are taken into consideration:

- Who acts ? What ACTORS (persons or collectivities) appear in the headlines ? What positions in the power structure do they hold ? What nationality do they belong to ? Are they of local, national or global importance ?
- What sort of processes are communicated ? The ACTIONS/PROCESSES are listed in a three point scale which runs from CONFLICT over NEUTRAL to COOPERATIVE actions. In connection with the ACTOR-dimension one can analyse images of actors and events related to them.
- What institutional realms (THEMATIC AREAS) are evoked ? What images are characteristic of what institutional areas ? Which themes are more personalized or more related to processes ? Which of them are dramatized (QUALIFIERS) ?
- What VALUES (negative or positive) are quoted or activated ? In relation to what type of actor or what thematic complex ? (The so-called 'Jeerformel' is a 'Massivformel' in an pragma-linguistic sense, aiming at in-group integration.

For the time being, the analysis concentrates on a special sector of these questions. It looks at the following problem : As regards the names of persons (as a partial operationalization of the ACTOR-dimension) what is the amount and quality of cross-cultural interrelation observed in the symbolic universes of sampled newspapers (their headlines) from different cultures (countries) and sub-cultures (national regions) ? A political indicator is expected to be developed out of the gathered data.

The coordination of the project rests with the Seminar für Sozialwissenschaften, Hamburg University. For the

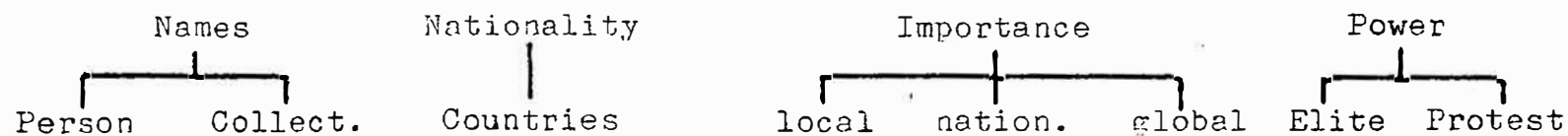
time being it is supported only by the infrastructure of the Institute. Contacts have been established with different researchers and it is likely that more collaborators will join in the project. Xeroxed papers are available. The Hamburg group is collecting headlines from the United Kingdom, the German Democratic Republic, from France, Spain and the German Federal Republic and will soon present first data. Text can be read automatically. A small multi-lingual dictionary has been built up. Data production is done by EVA (4).

Important areas of social indicator problems related to verbal interaction materials have not yet been discussed : What kind of content patterns are 'relevant' for social reporting ? How should a multi-dimensional indicator -and not only a category- be built ? By what method ? By what special procedures text indicators can be integrated into problem areas of the indicator approach ? What are the possibilities of text analysis for comparative social research ? We expect that working at the project will enable us to formulate problems more explicitly and to demonstrate some solutions.

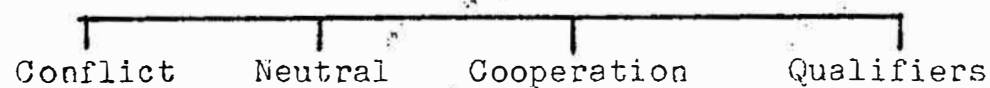
NOTES

- (1) See Ph.J.STONE et al., Newspaper Headlines: A Multinational Content Analysis Project on 'Textual' Indicators from Mass Media, SOCIAL SCIENCE INFORMATION 14 (1) pp. 112-116. For a survey of literature related to text indicators see A.DEICHSEL, Textindikatoren - Ein Literaturbericht, in: H.J.HOFFMANN-NOWOTNY (ed.), Politisches Klima und Planung, Frankfurt/New York 1977, pp. 99-116. For a detailed list of all correspondents and their research suggestions see A.DEICHSEL, Mass Communication: The Headline Project, in: A.DEICHSEL et al. (eds.), Maschinelle Inhaltsanalyse, Materialien I, Postprint of the meeting on Content Analysis at Hamburg, 11.-13.July 1975, Hamburg 1976, pp.78-89.
 - (2) As regards problems of sampling in content analysis research see now M.KOPS, Auswahlverfahren in der Inhaltsanalyse, Meisenheim 1977.
 - (3) See E.MOCHMANN, User's Guide to Content Analysis software, SOCIAL SCIENCE INFORMATION 14 (1) pp. 122-125. (*)
 - (4) The address of the Hamburg group: Seminar für Sozialwissenschaften, Universität Hamburg, Sedanstr. 19, D-2000 HAMBURG 13. The sample includes THE TIMES, The GUARDIAN, NEUES DEUTSCHLAND, LE MONDE, LE FIGARO, LA VANGUARDIA, ARRIBA, SÜDDEUTSCHE ZEITUNG, DIE WELT, RECKLINGHAUSER ZEITUNG, BAYRISCHE RUNDSCHAU. We took the central headline of the German and the most important two headlines of the other newspapers. The time period of observation runs from 1.10.76 to 31.3.77 for a first analysis. This sample has little theoretical value. It is constructed by accessibility reasons (as to the non-German units). Newspapers from West Germany are chosen from a larger sample of German newspaper.
- (*) See also the survey in: GMD/SIZSOZ-Expertisen, Ausgewählte Gebiete sozialwissenschaftlicher DV-Anwendung, Band 1, Linguistik, Inhaltsanalyse, Ökonometrie, Unterrichtsauswertung, Bericht Nr. 32 des Institut für Software-Technologie, Schloss Birlinghoven 1976.

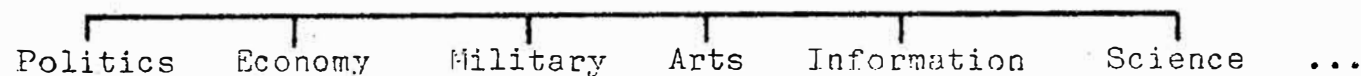
ACTORS



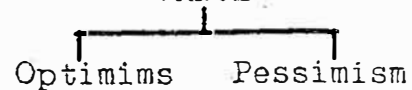
ACTIONS/PROCESSES



INSTITUTIONAL REALMS/THEMATIC AREAS



VALUE



TOPICS FOR WORK GROUP DISCUSSIONS/DECISIONS AT ANN ARBOR

CONCEPTS - VARIABLES

- What are 'relevant' indicators from mass media content to use in comparative social reporting (as a sector of comparative social research) ?
- What examples can be given for macroscopic content analysis and text indicator surveys ? (GERBNER, DEWEESSE, FLORA ...)
- Given the four dimensions of the Headline-Project, what value do they have for comparative reporting ? May they be the basis for developing a 'political indicator' ?
- What are the differences between a category in the ordinary sense and an indicator in the special sense of the social indicator research ?
- What examples can be cited for 'possible' indicators (as Sensationalism, Social Dynamics, Integration/Desintegration, International Relation, Concentration of Power, Mobility of Attention/Anomie, Ideology, Pessimism/Optimism ...) ? Which of them are feasible ?

MATERIAL TO EXAMINE

- What material of verbal interaction should be taken into consideration ? What sort of 'documents' are of value for further pilot-studies ?
- If one takes documents from different power strata how must be weighten the texts (for instance in EDUCATION : Texts from the government, regional bureuacracy, local directors, civil protest mouvements, counter cultures, newspapers to describe the symbolic oecology as regards a specific topic) ?

PROBLEMS OF OPERATIONALIZING

- What are the advantages/disadvantages of multi-lingual over mono-lingual analysis (translating all texts in one language) especially for comparative methodology (i.e. different extension of word-fields ...) ?

Examples for the two approaches ?

- What type of dictionary should be used for comparative purposes in the short run ? Analysis by a multi-lingual dictionary without disambiguation or coding by technically different dictionaries ? Or both ?
- How should be organized the crucial work of categorizing ? How to get more validity ?
- For which ~~special~~ special research questions the so-called empirical technics (WORDS) are preferable ? For what type of document ?

SAMPLING

- If one takes a specific research question (i.e. the one of the Headline-Project), what cultures/nations are of importance ?
- How should the sample of newspapers in a special country be built ?
- What parts of text should be sampled ? What headlines ?
- Are there theoretical considerations as regards the relation between the word in the headline and a specific word population (i.e. all headlines, newspapers words, written words, manipulated words in a given society) ?

CODING

- What software should be used ? With or without disambiguation ? A la QUESTER or à la INQUIRER ? How cooperate with manual coding ?

- Where should the coding be computed ?
Centralized or decentralized (i.e. at three places in West-Europe, East-Europe and USA) ?

DATA ANALYSIS

- How to build a text-indicator, conceived as a multi-dimensional index ? What appropriated technics exist to construct them out of word categories ?

INTERPRETATION

- What are cross-cultural areas of indicator research which can be completed by text indicators (i.e. Politics, Education, Health, International Relation, Development ...) ?
- How the intern-extern problems may be attacked ?
In what sense 'artefactual' and 'other' social indicators may be connected by social science theories ?

ORGANIZING FURTHER STEPS

- Shall we improvise further on at Hamburg ?
- How can we implement the results of discussions at Ann Arbor in the ongoing investigations ?