SAMPLING PROCESS-GENERATED DATA IN ORDER TO TRACE SOCIAL CHANGE

Nina Baur and Christian Lahusen[1]
Otto-Friedrich-Universität Bamberg, Germany

In this paper, we discuss sampling problems and strategies for process-generated data. For this purpose, we centre on the media discourse on unemployment as an example, here particularly on the news coverage by newspapers. This data provide us with a rich source of information because newspapers are an actor, arena and archive of public discourses at the same time. Depending on how the target population is defined, different sampling problems arise. Process-generated data are biased during data production because they are produced for other purposes and under other contexts than scientific research. Data preservation is biased, too, as data can be both destroyed and decay. Thus, humans have to actively preserve data for later use. Bias and the definition of target population reduce choice of sampling procedures. Specifically, random sampling usually is no suitable sampling technique. Instead, we suggest using sampling procedures usually applied when sampling qualitative data. We then suggest different strategies for applying these techniques for sampling newspaper data. Generally, sampling strategies are multi-staged. Researchers also have to decide how many newspapers, years, issues per newspaper and year and articles per issue are to be sampled. We finally demonstrate two contrasting ideal-typical strategies for handling this problem. These sampling strategies provide different solutions to the above stated problems. At the same time, they allow triangulating one sample with the other, thus controlling biases and overcoming limitations of a single sample to a certain extent.

Keywords
Process-Generated Data, Newspapers, Target Population, Cases, Sampling Bias, Multi-Staged Sampling, Adaptive Sampling, Most Different Cases Design, Longitudinal Research, Time Units, Time Span, Validity, Information Accessibility

1. INTRODUCTION

Empirical research in the social sciences is dependant on good data. This means also that valuable research questions often cannot be answered due to the lack of data and / or deficiencies of available data sets. One possibility to circumvent these problems is using process-generated data as an alternative form of data retrieval and generation. In this paper, we demonstrate typical sampling problems for this data type. In order to do this, we will focus on one specific type of process-generated date: reports on unemployment in German newspapers. We will first introduce the research question we use as an example. Then we will discuss why it is necessary to use process-generated data to address this question. We then discuss three problems that have to be solved in order to sample process-generated data: Researchers first have to

[1] E-mail: nina.baur@sowi.uni-bamberg.de; christian.lahusen@sowi.uni-bamberg.de

define the population. Second, they have to consider how data are biased during data production and selection. Using this information, they finally have to develop the actual sampling strategy.

2. EXAMPLE RESEARCH QUESTION: REPORTS ON UNEMPLOYMENT IN GERMAN NEWSPAPERS

This case of reports on unemployment in German newspapers illustrates nicely the advantages of process-generated data. As we will show in more detail later on, social scientists have developed a particular interest in public discourses about social problems (van Dijk, 1997). The development of public discourse on unemployment is one, yet, very suggestive example. In fact, since the 1970s, unemployment has been continually rising in most European countries. Thus, the last 30 years were characterised by heated discussions on how to reduce unemployment. During this time, mainstream academic discourse has shifted from (Neo-)Keynesian to Neo-Liberal arguments. Today, three argumentative structures compete: Both (Neo-)Keynesians and Neo-Liberals want to reduce unemployment and strengthen capitalism at the same time. While Neo-Liberals want to reduce social security, Neo-Keynesians try to save the welfare state. A third group of theorists seek an alternative to capitalism (Baur 2001).

While scholars have been analysing either *scientific* or *intellectual* discourse (Baur 2001), it is an intriguing to know how *public* discourse has changed over time: How important is the topic "unemployment" in public discourse? Who participates in public discourse on unemployment? Which arguments and solutions do discourse participants favour? Does public discourse refers to values and norms? And are public debates changing over time? We have tried to answer these questions in two related projects. The first project is part of an comparative project financed by the European Union. Christian Lahusen's research team analysed German discourse on unemployment in the newspaper "Die Süddeutsche Zeitung" from 1997 to 2002 (for the project design and codebook, see Giugni / Statham 2002; for first results see Baglioni / della Porta / Graziano 2004, Baum / Lahusen 2004, Chabanet / Fay 2004, Linders 2004, Statham 2004). In this paper, we will call this project "Project A". "Project B" was a local project conducted by a Nina Baur and Christian Lahusen together with a team of students. We analyzed discourse in ten German newspapers from 1964 to 2000.

3. WHY PROCESS-GENERATED DATA ARE NECESSARY FOR OUR RESEARCH QUESTION

Public discourses are a fairly new and interesting area of analysis. However, research on this topic is confronted by a serious problem: the lack of data. In fact, if – as in our examples – the research question demands tracing change over several decades, researchers need data from these periods. Many researchers face the problem that neither survey data nor qualitative interviews are available or adequate for their particular research question. Some of the reasons for this lack of data are (Baur 2004):

1) *New research questions may arise.* For example, until the 1970s, German employers actually had difficulties finding employees. Because of full employment, unemployment was no topic at all. In the following years, the "facts" about unemployment were discussed. Only in recent years, researchers became interested in public and scientific discourses on unemployment. As this topic had been regarded as mostly unimportant or uninteresting, no relevant data were collected. This will always happen: As society changes, social scientists' research questions will change.

2) *In the last decades, German sociologists focussed on individual change*, e. g. on individual attitudes to work, individual employment careers and / or their personal fear of becoming unemployed. In contrast, discourses are collective phenomena. The units of analysis in the sense of sampling are not "individual persons" but claims on unemployment (for a definition of "claims", see Giugni / Statham 2002). Put together, these claims form the discourse. In Germany, hardly any databases on discourses exist.
3) *For decades, German social scientists have also focussed on surveys*. Questionnaires cannot answer certain types of questions. We argued that discourses transcend the individual. Moreover, while issues and agendas might change more quickly, the argumentative or ideational basis of discourses evolves commonly over large time spans – years, sometimes decades. Here, we can speak of short and long cycles of changes. Thus, persons might simply not be aware of important characteristics of the discourse.
4) *Social scientists have continually improved their methods over the last twenty years:* New types of questionnaires, new analysis procedures and new design types have been developed. For example, event history analysis and sequence analysis are fairly new procedures. If researchers wish to benefit from these methodological improvements, they might not be able to use older data because these data might be appropriate only to old-fashioned methods. For example, event history analysis demands event data. In Germany, event data have only been widely collected since the end of the 1980s.
5) *For cross-cultural comparisons, respective data sets need to be available for all countries.* For example, in Project A, different countries' discourses on unemployment were to be compared. Even if there was a appropriate dataset for one country, it is highly unlikely that a similar dataset exists for the other countries, especially as these datasets not only need to be comparable in content. They also need to cover the same time span.

In our projects, lack of data could not be resolved by using retrospective interviews: We were interested in perception of discourses at the time. If we interviewed people today, several problems might arise: First, relevant participants in discourses might be dead or not traceable. Second, these participants might not remember all relevant details. Third, they would construct their version of the past from their present viewpoint. Relating to the discourse on unemployment, their opinion on things might have changed, they might not know the full story – which actors were successful and which were not. Consequently, people will tell a different story today than they would have told twenty years ago.

In cases such as these, an alternative is using process-generated data. Process-generated data are data not produced for scientific research. Instead, they are the result or by-product of social processes. Examples are newspaper articles, contracts, laws, speeches, records, files, protocols, diaries, personal notes, emails, letters, websites, databases, internet protocols, clothes, commodities, tools, furniture, architecture, landscapes, photographies, films, comics, paintings, sculptures, maps and so on. Thus, the array of documents, that can be used as a source of information for scientific research, is wide. They can already be standardized, semi-standardized or not at all standardized. Usually, the information they contain can be transferred to a database using qualitative and / or quantitative content analysis. Alternatively, process-generated data can be analysed using interpretative methods. In our example, we used newspaper articles, a source of process-generated data that so far has rarely been used (Müller 1996). We both transferred them to a data base and analysed them using hermeneutical methods (for details on mixing these methods, see Lahusen / Baur 2004).

However, before analysing process-generated data, researchers have to collect them. Usually, the amount of available process-generated data is so high, that they have to be sampled. In contrast to researchers using surveys and narrative interviews, researchers using newspapers

(and other kinds of process-generated data) so far rarely apply systematic sampling procedures (Lerg / Schmolke 1995). Sampling strategies determine if and how results of data analysis can be generalised. In order to determine if a sample of data is a reliable source of information, researchers have to address three problems:

1) *Researchers have to define the target population and the cases this population consists of* (Behnke et. al. 2004). As we will show in chapter 4, what is a fairly simple task for surveys, is much more complicated for process-generated data. The reason is multiperspectivity of process-generated data.
2) *If researchers want to test hypothesis or calculate confidence intervals, they need random samples. Thus, the sample should not be biased* (Behnke et. al. 2004). As we will show in chapter 5, process-generated are often biased during data production and data selection. If and how data are biased, depends on the data type and the target population.
3) *Using the information on target population and biases, researchers have to develop an actual sampling strategy*, that is, they have to select and find relevant cases. We will show in chapter 6, that usually a multi-staged sampling strategy is necessary for this. We will also show that different kinds of samples can be used to evaluate biases of other samples and thus make results more reliable.

4. TARGET POPULATION AND MULTIPERSPECTIVITY OF PROCESS-GENERATED DATA

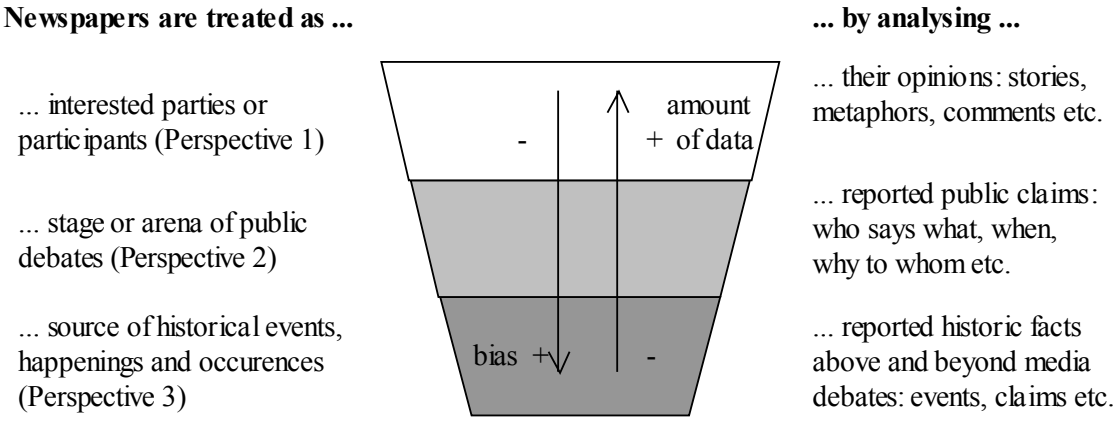
The first step in sampling is the definition of the target population (Behnke et. al. 2004). For surveys, this is usually relatively straightforward. For example, if researchers want to analyse Germans' attitudes on unemployment in 2004, the target population are (depending on how you define the term "German") either all ethnic Germans in 2004 (regardless where they live) or all persons living in Germany in 2004. A case is one person belonging to this population.

In contrast, defining the target population for process-generated data is more complicated. The reason resides in ambivalence of data – they can be read from a lot of different perspectives. For example, mass media evolved in modern society as an instrument to observe and report about reality: In Germany, in the 18th century historical science was divided into historical science and journalism. While the former were from now on responsible for tracing long-term development, the latter were responsible for writing "daily" history. Sharing the same roots, both occupations have been conceived as being bound to neutrality and seeking truth (Keppler 2000; on the history of German newspaper system see also: Bohrmann 1999; Kepplinger 1999a; Schütz 1999; Wilke 1999a). In terms of functionalism we can argue that the more societal actors and systems became dependant on mass media, the more did the value of newspapers consists in the production of reliable information. However, this is only part of the story (MacQuail, 2003). Mass media are also an actor or institution in their own right, with proper working routines and organizational needs. When we read newspapers or look at TV-news, we do not get an (unbiased) picture of reality, but rather journalists' highly selective view on social processes.

Both aspects are intertwined: Media are not (a mirror of) reality, although they are obviously a highly institutionalised form of describing reality. While theories of mass media have tended to arbitrarily concentrate on the one or the other aspect, we argue that this ambivalence is a specific quality of news coverage. Texts produced by media are an amalgamation and conflation of different facets of reality. Thus, researchers can read texts in multiple ways, depending on their specific perspective or research question. For example, newspapers can be read and analysed at least in three different ways (see figure 1):

In *Perspective 1* is to analyse mass media data in order to catch the story the journalist or the newspaper has to tell. In this case, *the journalist herself or the newspaper itself* is our object of

Figure 1: Possible Perspectives when Reading Newspaper Articles



research. We treat the media as an actor who is one of many actors participating in public discourses. Other actors are governments, interest groups, individuals etc. While newspaper articles might be filled with information about facts, events and claims of other actors, we would solely be interested in extracting the media's core message. Thus, we would focus primarily on the specific semantic organization of information, (implicit) comments and rhetoric devices (metaphors, catch words, examples etc) in order to capture what the meaning and message the newspaper or journalist ascribes to reality. In other words: Researchers analyse how journalists and papers report on what other actors did and how they behave in the discourse (Pietilä 1992). If researchers are interested in a journalist, the target population are all articles this journalist has ever written. If researchers are interested in a newspaper (for example "Die Süddeutsche Zeitung"), the target population are articles that have ever been published in the particular paper. If researchers are interested in more than one journalist or paper, the target population is what has ever been published by all respective journalists or papers. For all these examples, a case is one single article. If researchers are interested in Perspective 1, newspaper are a quite reliable source of information: As the discussion in chapter 5 will prove, there is no data production bias, although there still may be a data selection bias. Biased reflection of reality is not a methodological problem but rather the specific information we want to gather. At the same time, the amount of data is abundant. In fact, it is usually too abundant. As we will show in chapter 6, researchers usually can only choose a small sample of data for this case.

In *Perspective 2*, researchers conceive mass media as a public arena or stage (Gleich 1998; Schmitt-Beck / Pfetsch 1994). Some social scientists even argue that the mass media have become the most important turntable of the public space, of "Öffentlichkeit" in modern societies (Jansen / Ruberto / Münch 1997, Münch 1997a, 1997b). While face-to-face encounters and public meetings or events are a dominant form of establishing public spaces on the level of daily life, it is the media which provide a stage for public information and opinion-formation on the macro-level of complex societies. For Perspective 2, it not newspapers or journalists are of interest but the *public discourses* newspapers report on, although journalists may be part of this discourse (Kepplinger 1994, 1999b; Hoffmann / Sarcinelli 1999). We are not interested in deciphering the (implicit) opinion and story of the journalist. Instead, we centre our data retrieval on information about which actors have done or said what, when, why, and in view of whom. This enables us to reconstruct and understand the structure and dynamics of public discourses, as

they form within an important public arena (i.e., the mass mediated public). The target population are all actors participating in a particular public discourse, for discourse on unemployment. A single case is a claim made by one these actors. As newspapers are an important arena of public discourse, a lot of information on public discourses can be found in newspaper: Again, only a small sample can be analysed. However, data are never complete: Only a part of the overall discourse is reported on in newspaper. How big this window on the overall discourse is, depends on several factors: (Mass-mediated) public discourses are highly selective and / or exclusive. Not all actors and claims are able to enter the mass-mediated public sphere equally. Other actors exclude some possible participants from the discourse out, and media do not report evenly on all actors (Bright et. al. 1999). In fact, journalists today are the major gatekeeper in public discourse (Münch 1993, Rupp 1997, Wolff 2002). Different newspapers report differently on the same discourse. For example, German newspapers report more on German discourse on unemployment than British newspapers. German tabloids report less on this discourse than high-end newspapers. While German newspapers discuss unemployment a lot, they generally do not talk about other topics, for example on how to grow cactuses. The perspective on gets on public discourses by analysing newspapers is thus broken by the particular newspaper's view on this discourse. Thus, researchers using Perspective 2 have to handle both data production and data selection biases. Still, data are only partly biased, as newspapers usually are the main arena of this discourse. In order to perceive the full discourse, researchers would have to analyse not only all newspapers but also other media (e.g. TV, radio, internet, books etc.).

In contrast, *Perspective 3* regards newspapers as a medium and / or as a (biased) window to *reality itself*. Press coverage is seen as an archive for historical facts (Franzosi 1987). In the case of political contentions about unemployment policies, for instance, we would assume that the press gives us a more or less restraint testimony of these conflicts – trusting that the media presents us at least part of the story. The target population are all facts, events, occurrences and happenings regarding these conflicts. Single facts, events, occurrences and happenings are the cases. How much information on actual social processes (other than public discourses) researchers can draw from media data depends on the type of discourse: As we will show later on, one could receive quite reliable information on German official unemployment figures or strikes from German newspapers. It is a lot harder to find out how the unemployed themselves live and behave. Still, this information is at least a little reliable, as unemployment is an important topic for Germans. In contrast, it would be probably be impossible to use this data source to find out if there are import regulations on cactuses and how they were determined. All in all, *Perspective 3* is the perspective where amount of data is smallest and where data at the same time most biased: Reality is broken twice – through public discourse's and through journalists' perspectives. As researchers are interested in facts, they should triangulate data (e. g., reports from other newspapers, interviews, documents; on triangulation see Seale 1999; Flick 2000).

In summary, newspaper articles reflect multiperspectivity of reality: They can be read in different ways, drawing different kinds of information from them. This ambivalence does not apply uniformly to all media types. In practice, communications studies have centred on TV-channels as actors of public discourses because hey have been interested in unveiling the impact of news coverage on public perceptions and opinions (e.g., Brooks 2004; Chiricos / Padgett / Gertz 2000; Tudor 1992). TV is a better candidate for this kind of research because the selectivity and construction-process is much more evident in the case of television than in the case of newspapers. It is no surprise that newspapers are therefore used more recurrently when researchers analyse public discourses and / or are looking for archives of empirical data. In this paper, however, we argue that newspapers mirror the above mentioned ambivalence or multiperspectivity most perfectly because there are actors, arenas and archives at the same time.

Depending on the research question and perspective on the data, different kinds of target populations and cases have to be defined. Different perspectives in turn face different kinds of

biases, have to handle different amounts of data and use different sampling strategies. This ambivalence does not necessarily create problems for empirical research, as long as we know what kind of information we can and want to ask from data, and as long as we develop strategies to control the inherent bias of this medium. Hence, from a methodological point of view, the latter aspect needs more consideration before sampling strategies can be developed.

5. RANDOM SAMPLES AND BIAS

The whole idea of sampling is to create a sample of data that allows to generalise findings after data analysis. One typical way of generalising is using hypothesis testing or confidence intervals. Both statistical procedures require random samples. In other words: If researchers want to apply inferential statistics, the sample may not be biased (Gigerenzer 1999; Mayer 1998; Behnke et. al. 2004). For surveys, reasons for bias are usually bad research designs, nonresponse and missing values (Schnell 1986, 1997). In panels, panel mortality and spell effects may additionally bias data (Blossfeld et. al. 1986; Steinhage 2000). Additionally, the target population may change over time (Abbott 2000, Baur 2004). Most of these problems can be handled or at least minimised by a careful research design. Therefore, survey can but do not necessarily have to be biased.

The situation is different when using process-generated data: Process-generated data are almost always biased. This means, the target population and the frame population differ. The frame population over covers some types of cases of the target population, other types of cases are under covered or even completely absent from the data (Behnke et. al. 2004). Two intertwining processes might influence the bias: As researchers cannot control the process of data production, data usually are already biased during production. Moreover, this bias will accumulate over time: more and more original data might decay or will be destroyed deliberately. This process, too, is biased, as humans have to actively want to preserve data available for later use. The researcher can influence neither data production nor data selection process. However, both processes influence what data are available at all. Furthermore, these biases depend on the research perspective on newspaper data and on the target population: As long as researchers analyse newspapers as participants of public discourse (Perspective 1), data production bias is not a real problem but rather a peculiarity of this data type. As soon as we treat newspapers as a public arena or historical archives (Perspectives 2 and 3), data production bias becomes pertinent. In addition, researchers have to handle data selection biases for all three perspectives. Researchers cannot change this situation. However, learning from historians, social scientists might evaluate biases by qualitatively analysing production and selection contexts. Before demonstrating these possibilities by using our data as an example, we will discuss the reasons for data production and data selection biases in more details.

5.1 Biased Data Production Process

Data are supposed to document events and processes. Because so much is happening at the same time, a full account of reality is impossible. This is as true for both process-generated data and for survey data. However, in contrast to surveys, researchers cannot control the process of data production when using process-generated data because the latter have been generated for other, practical purposes, e.g., to inform people on a daily basis on local, national or international news. Thus process-generated data are usually already biased during data production.

But *how* are they biased? This depends on the particular purpose, format and institutional embeddedness of the pertinent data type. In addition, all three elements may change over time.

Researchers need to know this context well in order to be able to assess how the bias works. They might possess this knowledge because they are doing research on a period and topic of their own culture well-known to them. If not, it is necessary to obtain this knowledge by using qualitative and historical research methods. When using Perspective 1 for analysing newspaper articles on unemployment, there is no data production bias, as the newspapers themselves are the focus of interest. For Perspectives 2 and 3, discourses and social processes are reflected in the newspapers. This means, there usually is a data production bias. To find out, how this data production bias works, one can first analyse newspapers using Perspective 1, or one can use other data types. In the case of the newspaper data, the following elements have a strong impact on the production bias:

1) *Purposes:* Journalists are human, and as all humans they perceive their environment selectively. This selectivity reflects theirs specific biography and socialisation, their ontology and values. Data production bias increases if journalists are not aware of their specific perspective (Kepplinger 1994, 1999b). Although this kind of bias can never be abolished, the German news system tries to decrease it: In general, German journalists pursue the professional goal of informing the public, contributing to the formation of public opinions and controlling state power, e.g., by means of "investigative journalism" ("investigativer Journalismus"). This professional self-concept applies particularly to the press, which stresses its public mandate quite clearly (Keppler 2000; MacDevitt 2003). Rules and checks have been instituted both on the level of daily working routines (e.g., the separation of information and commentary in news coverage) and on the level of institutional controls (e.g., liabilities of newspaper publishers) in order to safeguard that these purposes are met effectively (Pöttker, 2002). At the same time, the media system inherently tends to increase data production bias: German mass media are increasingly dominated by economic rationality (Zerdick 1994, Altmeppen 2000, Küng 2001, Oberst-Hundt / Oberst 2001): Globalisation (Kleinsteuber / Thomaß 1996, Hagen 1996, Wittenzellner (ed.) 2000) and concentration (Hell-Berlin 1997, Groß 1996, Jarren 1996, Röper / Pätzold 1993) increase competition within global and German mass media. Due to economic pressure, professional ideals very often are forsaken. The boundaries between information and entertainment (Wolf 1999, Zerdick 1994) as well as the boundaries between journalism and public relations are increasingly blurring (Jarren / Meier 2002, Trappel / Meier / Schrape / Wölk 2002, Kiefer 2000). The few studies that exist on this topic suggest that journalists all over the world are far from the ideal of investigative journalism because they do not know enough about the topic, they do not work continually on the same topic (to build up this knowledge) and because they do not have the time for their own research (Bow 1980, Kuenheim 1996).
2) *Formats:* Newspapers conform to various formats. On the level of the individual article we know that information is selected according to specific rules (e.g., news values, see Galtung & Ruge, 1973) and that articles follow a particular pattern and structure (e.g., summary, main-event, background, comments; see van Dijk, 1988). On the level of individual newspaper issues we are dealing with different sections (e.g., national and international news, business, culture, sports), which have quite different agendas and a particular approach towards news coverage. This means, for instance, that unemployment will not rank high in sport or culture sections. Additionally, reports on unemployment will have an entirely different story to tell in these sections, when compared to the pages covering political or business news. On the level of newspapers we may distinguish between prestige newspapers and tabloids, daily and weekly, local and national newspapers, papers with a broad scope of topics and highly specialized newspapers, amongst others (Kepplinger 1994; Wilke 1999b; see also table 1). Newspapers usually report more detailed on events and discourses that take place in or near the region they address (Oliver / Maney 2000). For example, „Der Fränkische Tag" and „Die Nürnberger Nachrichten" are printed in Franken. These regional papers

thus report more on events in Bavaria, while „Die Stuttgarter Zeitung“ addresses and reports more on events in Baden-Württemberg. Even national newspapers are regionally biased. For example, “Der Spiegel” reports disproportionally on events in Hamburg, where it is printed. Formats and content vary regionally, too. For example, in “Die Stuttgarter Zeitung” a lot of information on unemployment can be found in the section “Business News”, while the same topic is addressed in the section “Politics” in “Die Süddeutsche Zeitung. Also, newspapers report in a more detailed fashion on events that are defined important concerning the format while they might blend out other topics completely. For example , the topic “unemployment” is highly relevant for modern German newspapers, while “cactuses” are not. One might find the latter information more easily in a gardening journal. On all these levels, formats and purposes are strongly interrelated. For example, the editorial definition of events influences strongly the way in which viewpoints are presented (van den Berg et. al. 1992). Moreover, the interaction between formats, purpose and content changes over time. As soon as we are acquainted with mass media, we know, for instance, which article, section or newspaper we have to read when we are interested more in information, opinion or entertainment.

3) *Institutional embeddedness*: On a more general level, the production-based bias is determined also by the environment of mass media. First, newspapers have to conform to legal guidelines of what and how to report (e.g., restrictions on pornography, respect of privacy). Second, most media (TV senders, radio stations, but primarily newspapers) have quite stable political allegiances (see table 1). Moreover, the fact that newspapers depend on valuable sources (in most of the cases within government, parties and interest groups) brings about political loyalties or solidarities. Finally, mass media depend also from markets, in the sense that they are interested in securing or expanding the range of advertisers and readers (On the way German readers use newspapers and newspaper contents see: Kiefer 1998; Schulz 1999). All these institutional environments will have an impact on news reporting, e.g., by privileging certain topics, positions, formats or purposes to the detriment of others. Again, institutional factors are intertwined with journalists’ purposes and formats: Due to increasing competitive pressure, journalists today have less time for a single article. They increasingly use companies’ and political parties’ press releases instead of their own research for articles. Thus, they tend to be biased towards specific actors. Also, institutions have their own rhythms: Election cycles, typical dates for press releases, parliaments’ sessions and so on. This means, there are more news on some topics at specific dates, less on

Table 1: Some German Newspapers

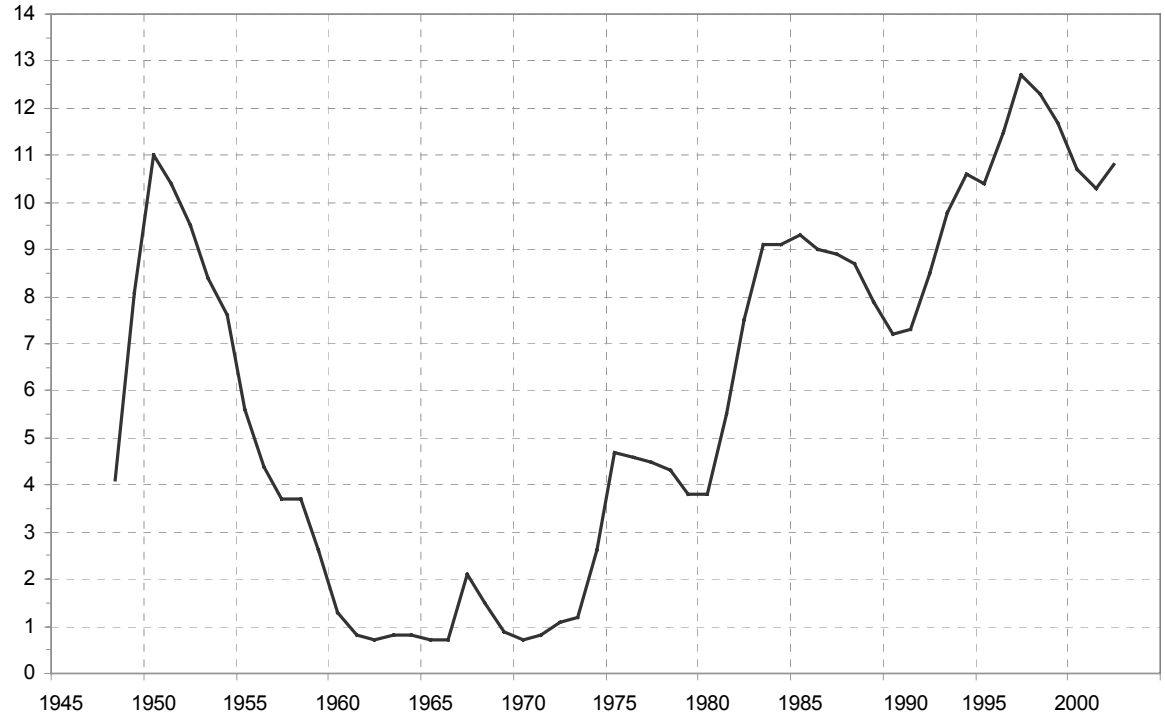
Newspaper	Rhythm of Publishing	Readers	Political Alignment
Der Spiegel	Weekly	National	Liberal Left
Die Zeit	Weekly	National	Liberal Left
Die Frankfurter Rundschau (FR)	Daily	National	Liberal Left
Die Süddeutsche Zeitung (SZ)	Daily	National	Liberal Left
Die Bild-Zeitung	Daily	National	Conservative
Die Frankfurter Allgemeine Zeitung (FAZ)	Daily	National	Conservative
Die Welt	Daily	National	Conservative
Fränkischer Tag (FT)	Daily	Regional	Conservative
Nürnberger Nachrichten	Daily	Regional	Social Democratic
Stuttgarter Zeitung	Daily	Regional	Conservative

others. This affects reports on other topics, as the amount of slots in a given newspaper is limited (Oliver / Maney 2000). "Unemployment" is a high priority topic in Germany, thus "stealing" slots from other news. However, there are exceptions: Catastrophes, wars, elections and other news with a high news value crowd out other news (Schmitt-Beck / Pfetsch 1994; Oliver / Myers 1998). They might also push "unemployment" to second place.

In summary, newspapers cover issues differently depending on time of occurrence, location of occurrence and the relevance of the issue for the social context (Hocke 1998; McCarthy et. al. 1996; Oliver / Maney 2000). In order to assess selectivity, researchers therefore have to analyse the social context. For example, we consider the following types of context information as especially important for the amount of news coverage (and therefore for production bias) on the German discourse on unemployment:

1) *The Development of Unemployment:* Discourses on unemployment are not independent from the actual development of unemployment (see also figure 2). In Germany, the 1950s and 1960s had been earmarked by full employment. Actually, there was a lack of labour. Since the middle of the 1970s – starting with the oil shock – unemployment has been continually rising. After unification in 1990, unemployment figures jumped (mainly due to high unemployment figures in East Germany). This means, that in the 1950s and 1960s public

Figure 2: Official Unemployment Rate for Germany (in %)



Definition of unemployment and / or measurement methods were changed in 1966, 1985, 1987, 1989 and 2004. The figures also measure unemployment rates for different regions: Until 1949 for West-Germany without Saarland and West Berlin; from 1950 to 1958 West-Germany without Saarland; from 1959 to 1990 for West Germany; since 1991 for West and East Germany

Source: Bundesagentur für Arbeit, 24.05.2004, http://www1.arbeitsamt.de/hst/services/statistik/aktuell/iiia4/zr_alo_qu_ab_1948d.xls

discourse centred on the lack of labour. One can draw almost no information on unemployment from German newspapers for this time span. In contrast, unemployment has been a central topic in German newspapers ever since the 1980s.

2) *Historical Legacies:* In Germany, unemployment is a sensitive topic: During the 1920s and 1930s, social climate was very instable. Then unemployment rose rapidly during the recession in the 1930s. This combination made it possible for Hitler to win the elections. Thus unemployment is usually conceived as an immediate cause of destabilising democracy. This was especially important in 1997, when unemployment figures topped the unemployment figures of the 1930s for the first time since 1949 (Baum / Lahusen 2004). Additionally, unemployment receives always high coverage in comparison to other topics.
3) *Political System:* The political system also influences the way public discourses work. For example, it is likely that election years and election campaigns influence these debates, both in its intensity and internal structure. As unemployment is considered such an important topic, elections mean that there is even more news on unemployment, while other topics are crowded out. However, discourse's character changes during this time: Politicians are even more dominant in the discourse than usually (Schmitt-Beck / Pfetsch 1994). During election times, discourse is centres on staged events ("Pseudoereignisse") and on seeking culprits (for high unemployment) than on finding manageable solutions (Schmitt-Beck / Pfetsch 1994). Public discourses do not only reflect election cycles but also the federal structure of German polity. This means, for instance, that regional papers (or regional sections in national papers) become an important source of information when analysing regional debates. German news coverage is also known to be government centred. We thus need to know the succession of parties in government in order to estimate the effect of this bias on the inclusion or exclusion of political parties from news coverage (see table 2). Additionally, it is important to know that labour market policies are institutionalised in a neo-corporatist way: Trade union associations and employers association decide on many areas of labour

Table 2: German Governments since 1949

Chancellor	Governing Period	Governing Parties	Elections
Konrad Adenauer	September 1949 – October 1963	Conservatives (CDU / CSU) & Liberals (FDP)	14.08.1949
			06.09.1953
			15.09.1957
Ludwig Erhardt	October 1963 – December 1966	Conservatives (CDU / CSU) & Liberals (FDP)*	17.09.1961
			19.09.1965
Kurt Georg Kiesinger	December 1966 – October 1969	Conservatives (CDU / CSU) & Social Democrats (SPD)	28.09.1969
Will Brandt	October 1969 – May 1974	Social Democrats (SPD) & Liberals (FDP)	19.11.1972
Helmut Schmidt	May 1974 – October 1982	Social Democrats (SPD) & Liberals (FDP)	03.10.1976
			05.10.1980
Helmut Kohl	October 1982 – October 1998	Conservatives (CDU / CSU) & Liberals (FDP)	06.03.1983
			25.01.1987
			02.12.1990
			16.10.1994
Gerhard Schröder	Oktober 1998 – ?	Social Democrats (SPD) & Green Party (Bündnis 90 / Die Grünen)	27.09.1998
			22.09.2002

* Not all the time.

market policies. Different newspapers and different newspaper sections will provide us with quite different information on these organizations. For instance, business sections and business newspapers usually favour employers' organisations (Gesterkamp 1993). Finally, German welfare associations are not represented in labour market policies, while being heavily involved in social policies and, thus, in the practical work with the unemployed. If we want to reflect the public discourses and contentions in this wider spectrum, we would therefore need to select articles both on unemployment and labour market issues and on social security.

4) *Social Security Systems:* Actors will not talk about things they take for granted, only about issues debated. What is taken for granted and what is debated depends on the specifics of the social security system, as every welfare state removes certain social problems but is also a source for new conflicts. In other words: Just because people do not talk about certain problems, this does not mean, they do not exist. On the other hand, certain problems might be framed in a different context. Since the 1920s, Germany has had a extensive system of unemployment benefits. The national unemployment agency (formerly "Bundesanstalt für Arbeit", now "Bundesagentur für Arbeit" (= "BA")) also collects statistics on the development of employment. Germans thus conceive a person as "unemployed" if they are reported to be unemployed at the BA. It does not matter if they wish to work or not (Baur 2001). Thus, labour market difficulties of a lot of population groups are not debated as problems of unemployment. For example, Germany is a Conservative Welfare Regime (Esping-Andersen 1990) with a Strong Male Breadwinner Model (Ostner (1995). If women are laid off, very often they drop out of the labour market completely. Unemployment is considered mainly a male problem – by both men *and* women. Young people either go to university or they participate in the well-developed system of occupational training ("Ausbildungssystem"). Actors therefore never talk about youth unemployment but about a lack of trainee slots ("Ausbildungsplätze"). Unemployment security only covers certain occupational groups: people employed by companies ("abhängig Beschäftigte"). State employees ("Beamte") have a guaranteed life-time employment. The self-employed ("Selbständige") are not part of the social security system. Therefore, neither group is talked about in the context of unemployment. The same is true for ethnic minorities, but for different reasons: There is no separate employment statistic on ethnic minorities. Germany's citizenship is based on *ius sanguinis*. Therefore ethnic Germans – regardless if they were born in Germany or immigrated – are considered as Germans (German Information Center 1995). Persons seeking an asylum are not allowed to work. If they are discussed at all, it is in the context of moonlighting. In the 1960s Germany invited immigrants from Southern Europe (especially Turkey, Spain and Italy) to work in Germany as "guest workers" ("Gastarbeiter"). The idea was to send them back once they were not needed any more. Of course this did not work Today, the guest workers are either completely integrated in the labour market (having paid into the social security systems and thus being eligible to benefits) or have returned to their country of origin (Castles / Kosack 1985).. All this information is important for deciding how to newspaper data are biased. For instance, if researchers want to analyse discourses on youth unemployment in Germany, they would have to consider articles dealing with the training system, while this might not be necessary in other countries.

5) *Time Span of Important Discourse Changes:* We have argued that public discourses evolve according to short-term and long-term cycles: while news coverage changes quickly, when issues and agendas are concerned, this is not true when looking at underlying ideas, ideologies or world-views. For instance, academic discourse on unemployment shifted at the beginning of the 1970s from discussions about the distribution of labour to debates about the preventive avoidance of unemployment. Till the mid-1980s, academic discourse was dominated by a labour-friendly, Keynesian position. Since then it has been shifting to a em-

ployer-friendly, neo-liberal position (Baur 2001). It is likely that media discourse shifted correspondingly. In case we want to grasp these shifts, we would need to generate a sample of data that covers this longer time span.

These observations illustrate that mass media data are highly biased, if researchers use Perspectives 2 and 3, for instance, when dealing with the public discourse on unemployment. This means that it is impossible to draw random samples for research questions using these Perspectives (although one might draw random samples for Perspective 1). However, this problem does not discredit process-generated data as source of information. Instead, researchers can use alternative sampling strategies that are typically used in qualitative research (for an overview see Creswell 1998). Examples are the selection of typical cases, the most different cases design and the most similar cases design (Behnke et. al. 2004). We will demonstrate this in more detail in chapter 6.

In order to choose the appropriate sampling strategy, it is necessary to know how bias works for the specific research question. It has become evident that these biases are not randomly generated but instead highly patterned and structured. Media sociology has illustrated quite clearly the rules according to which the media generate their products. Hence, we can anticipate these biases in order to depict how mass mediated public discourses (Perspective 2) work. On this level of analysis, we can use the knowledge about the production bias in order to explain the particular structure and dynamic of public discourses in one important respect. If we are interested in using newspapers as a source of historical data (Perspective 3), newspaper data should not be the only source of information. Researchers can either triangulate newspaper data with other data types, e. g. interviews, books, pictures, diaries etc., or they can triangulate different types of newspaper samples: If it is true that mass media is an highly selective arena, then we can assume that we are dealing with a multiplicity of arenas. They all share the fact of being highly selective, but they all have their own specific bias: leftist and rights papers capture a different facet of the political landscape and a distinct (biased) window to reality. The same is true for national versus local newspapers, prestige press versus tabloids, business versus political sections of the same newspaper and so on.

5.2 Biased Data Selection Process

So far, we have discussed how data production biases processed generated data. When using process generated data to trace social change, researchers face an additional problem: Data have been produced at earlier points in time than they are analysed. During the time that has passed between data production, data preparation and data analysis, data can be lost. What seems harmless at first sight may seriously impair research results as data selection is biased, too.

First of all, human beings might deliberately destroy data because they do not want the information being kept for former generations. This happens very often to documents. A recent example is the German governmental change in 1998. When the Social Democratics took over the “Bundeskanzleramt” (the Chancellor’s personal administration), they discovered, that obviously important files were missing. Until today we know that there is something missing. However, we do not know exactly what files are missing and what information they contained. It is very unlikely that random data were destroyed. Probably, the missing data contained some information unfavourable to the former government.

Another reason why data might be destroyed is to make room for other data: Newspapers are published daily. They soon pile up and occupy a lot of storage room. Individual readers and researchers might decide therefore to throw away their personal archives. Even public libraries,

archives or publishers might choose to do the same if data have not been used a lot or are considered less valuable at that period of time. Moreover, data might get lost as a consequence of the bankruptcy of publishers or because archives are closed down and/or assigned to other entities. Finally, natural disasters or social turmoils might have their share in the destruction of data. In the case of newspapers, data destruction is generally no problem: Newspapers usually do not contain secret but public information. In addition, so many copies are printed and distributed, that it is very unlikely that *all* of them were destroyed.

A second factor that leads to a data selection bias is the decay of original data. For example, newspapers are usually printed on paper. Paper may burn. It may get wet and rot. The ink may dissolve or eat the paper, and so on. Therefore, human beings have to take active measures in order to prevent data from deterioration. They only take such measures for things they think important. Here again, we could assume that the danger of decay might not be a severe one, considering the fact that newspapers are mass products present in a number of different archives. However, while we have no empirical evidence, it is to be assumed that the danger of data destruction and decay affects less prestigious and more short-lived newspapers. We would expect that highly recognized newspaper will hardly disappear from human memory, while this might not be the case with ephemeral papers or tabloids.

However, newspapers are a commodity for everyday use. Therefore, not *all* but *most* copies of newspapers probably have been destroyed. Thus, the difficulty with newspapers, is finding the remaining copies. Researchers can rely on a variety of different sources. In most European countries, for instance, newspapers have been archived for at least 100 years either by the publishing companies themselves or by public agencies. For example, in Germany, each "Bundesland" has a central library or archive (the "Staatsbibliotheken" and "Staatsarchive"). These archives collect one copy of everything that has been printed in the respective region. Important German newspapers are also archived in University libraries and public libraries. Some archives and libraries collect newspapers or articles on certain topics as well. For example, staff of the "Hamburger Weltwirtschaftsarchiv" (HWWA) have been reading several hundred German newspapers every day since the 1970s. From these papers, they have been selecting all articles on companies or industries. Thus, for each company, a file exists that contains (almost) all articles that have been written on this company in Germany since 1970. For a small fee, researchers can copy these articles. In recent years, major newspaper producers have started digitalising their newspapers. Search machines can be used to find articles quickly. The articles then can be easily either printed or imported into CAQDAS. This sampling method was used for Project A. However, only a minority of newspapers publishes their articles on CD-ROM. In addition, so far only data since the beginning or middle of the 1990s are available. Thus, for Project B, we had to draw on libraries or microfilms in order to obtain the relevant information.

Although newspapers can be stored in multiple formats and locations, this does not have to be the case. Prestigious papers are present at various locations and archives. In fact, many public libraries and archives only collect major and some local newspapers. Moreover, these papers are available in various forms (paper issues, microfilms, CDs), while tabloids are less as accessible in terms of a variety of archives and well developed electronic data-bases. For example, we had to rely on a mixed sampling strategy for Project B: The major newspapers ("Die Zeit", "Der Spiegel", "Die Süddeutsche", "FAZ", "FR", "Die Welt") were available either in the Bamberg University Library or in the "Staatsarchiv" which is also situated in Bamberg. For the local newspapers, we actually had to drive to the company archives, which is one of the reasons we selected *these* newspapers: "Fränkischer Tag" is situated in Bamberg, "Nürnberger" in Nuremberg and "Stuttgarter Zeitung" in Stuttgart. These newspapers thus were easily accessible. We had problems obtaining relevant articles for "Die Bild-Zeitung": Being a tabloid and being printed in Hamburg, "Bild" is neither archived in any place nor close to Bamberg, and no electronic data base is available.

Accessibility is certainly not a severe problem, however, it does constitute an implicit selection bias. Researchers have to know the social context well in order to know *where* process-generated data are stored, how to find them and how to gain access to them. If there are several locations, usually the one closest to the researchers office is the one to be preferred in order to cut down sampling costs.

In summary, process-generated data can suffer from a data production bias and from a data selection bias. For newspapers, data selection is usually no problem, although researchers need to know the social context well in order to find respective data sources. For Perspectives 2 and 3, researchers additionally face data production bias. Therefore, it does not make sense to draw random sample from this data type. However, knowing how the bias works can help developing alternative sampling strategies.

6. STRATEGIC CHOICES FOR SAMPLING STRATEGIES

After having defined the target population and after assessing data production and selection bias, researchers have to develop the actual sampling strategy: Our indications have illustrated that mass media provide data in abundance. Not only are there several hundred German newspapers, there are also multiple other media: TV, radio, books and so on. Hence, in the case of public discourses the amount of process generated data is so large that researchers have to sample them.

These three steps are not independent but intertwined: Which sampling strategy is best suited depends on the specific research question. The particular perspective on newspaper articles influences both the definition of target population and cases and how much data are biased. When using newspaper data for Perspective 1, bias is generally not problem. Therefore, researchers can draw multi-staged random samples for this perspective. The procedure is the same as for surveys (e. g. Cochran 1972). The only difference is that articles are sampled instead of persons. Of course, researchers can also use the alternative sampling strategies we demonstrate below. In fact, if only a small number of cases can be sampled, this is advisable.

When using newspaper data for Perspectives 2 or 3, data are always biased. In this case, random samples do not make sense: Bias either over layers random error or completely render inferential statistics impossible (Gigerenzer 1999; Mayer 1998; Behnke et. al. 2004). Thus, on the one hand bias limits the choice in sampling strategies. On the other hand, knowing how the bias works can help developing alternative sampling strategies. We suggest using sampling strategies typical for sampling qualitative and historical data (for an overview see Creswell 1998; Behnke et. al. 2004).

What has to be taken into account for sampling? First, the unit of analysis is either a single article (Perspective 1) or an event or claim stated in this article (Perspectives 2 and 3). In order to find the units of analysis, researchers first have to select relevant newspapers and relevant issues . Thus, generally a multi-staged sampling strategy is required for sampling process-generated context: Researchers first have to decide which newspapers to analyse (Stage 1). They then have to find out where these newspapers are stored and select the issues to be analysed from these newspapers (Stage 2). From each issue, relevant articles have to be chosen (Stage 3). For Perspectives 2 and 3, events, facts and claims have to identified using these articles (Stage 4).

Second, in addition to defining the target population and assessing bias, researchers have to decide how they want to locate their sample on the following dimensions:

1) *the number of newspapers to be sampled*;
2) *the time-span, i.e. the number of years to be sampled*;
3) *the number of issues to be sampled per newspaper and year*;

4) *the number of articles to be sampled per issue*;
5) (only for Perspectives 2 and 3) *the number of claims or events per article*.

For each of these dimensions, the strata goes from a single-case design (one newspaper; one year; one issue; one article; one claim or event) to sampling the whole population (all German newspapers; all years of interest to the research question; all issues that have been published by selected newspapers in selected years; all articles in selected issues; all claims or events cited in selected articles). The result is a four-dimensional space of sampling possibilities. The extremes in this space would be a single claim, event or article versus all claim, events or articles that have ever been written (about) in Germany.

While being aware of these extremes is helpful for making strategic choices on sampling, neither extreme is suitable as an actual sampling strategy: One the one hand, one cannot generalise from a single claim, event or article. On the other hand, researchers can only spend limited time and money on a single project. Usually, a maximum number of articles, claims or events that can be analysed is given. This maximum number not only depends on funding but also on the planned analysis strategy: It is lower for qualitative methods than for quantitative methods. Given this maximum number of articles to be sampled, researchers face a trade-off: Going up on one dimension means going down on another.

Researchers cannot cut down dimensions without a price: Data become biased in the sense that data are not suited for answering certain kinds of questions. However, in contrast to the bias we talked about earlier, researchers can influence the bias produced when cutting down dimensionality. The trick is to reduce dimensions that are not needed for answering the research question. In other words: It does not matter of the sample is not suited to answer certain kinds of questions as long as these are not the questions researchers want to answer and as long as researchers are aware that these questions cannot be answered using the particular sample. For different kinds of research questions, high scores on different dimensions are important. Using the German discourse on unemployment as an example, we will discuss how researchers can decide how to reduce dimensionality. Note that there is no general solution for this problem. It can only be solved for specific research questions. Researchers need a definition of the target population. They also need to know the social context and how bias works.

6.1 Number of Newspapers

The number of newspapers has to be high, if researchers are interested in different position of different newspapers (Perspective 1) or in a wide spectrum of arenas, orientations and positions (Perspective 2). If historical facts are to be reconstructed (Perspective 3) using newspapers articles, it is also important to sample as newspapers as possible, as this allows researchers to triangulate different newspaper reports. If one of these aspects are important, it is sensible to make use of a most-different cases design.

For example, for analysing discourses on unemployment, it is sensible to sample newspapers that differ as much as possible on the following dimensions: (a) location of production; (b) regional vs. national scope; (c) publication cycle; and (d) political orientation. We used this sampling strategy in Project B, selecting from the papers in table 1. Note that we decided to concentrate on general newspapers, dropping out TV, the radio, the internet, books and special interest newspapers and journals from our frame of analysis.

In contrast, the team for Project A was not interested in variety of discourses but rather in "typical" discourse. Thus, the research team chose only one single paper: "Die Süddeutsche Zeitung" (SZ). The decision to use one newspaper reduces the spectrum of news coverage and increases thus the production bias. The SZ is one of the national prestige newspapers, and thus

particularly interested in covering national political news. It is considered as a moderately liberal newspaper, and thus represents the political mainstream with a weak leftist turn (see also Zakrzewski (1995)). This orientation seemed to provide us with a bias that fits well the research priorities of the project: The SZ is strongly focussed on the core policy domain (government and opposition, social partners, experts), while having a certain affinity to leftist organizations (e.g., the unions, welfare organizations, unemployed initiatives). As our project was interested in describing the fate of the unemployed and their organizations within the public discourse on unemployment, this choice seemed to be perfectly justified. Finally, the SZ has strong regional roots, as all German national newspapers have. By choosing the SZ, we opened a window to the federal structure of German polity, in this case to Bavaria, which plays a crucial role in German politics. Bavaria plays the role of an unofficial, Christian Democratic counter government to the Social Democrats, which are in power in Berlin since 1998. In spite of these regional roots, we checked that news coverage is less regionally biased than within other national newspapers, thus providing a balanced relation between national and regional orientations is generally considered as one of the most neutral national papers. This allowed increasing scores on the other dimensions for project A.

6.2 Time-Span and Number of Years within this Time-Span

If researchers want to reconstruct historical facts (Perspective 3), time-spans to be analysed should generally be short because both the number of newspapers and the number of issues per newspaper and year has to be high. Newspaper additionally should be triangulated with other data types. Thus, analysis is very time-consuming. There is one exception: If archives exist that collect articles on specific papers (e. g. the HWWA for German business news), data collection and analysis can be manageable even if the researchers analyse large time-spans.

For Perspectives 1 and 2, researchers have to decide if they are more interested in short-term or long-term discourse cycles. While news coverage changes quickly, when issues and agendas are concerned, this is not true when looking at underlying ideas, ideologies or world-views. When analysing the former, we would decide to investigate a shorter period of time in more depth, while doing the opposite when studying the development of ideas or ideologies. In both cases, researchers need to decide about time spans on the basis of assumptions about issue- and ideological cycles.

For example, using the context information discussed in chapter 5, we would assume that discourse on unemployment shifted at the beginning of the 1970s from discussions about the distribution of labour to debates about the preventive avoidance of unemployment. Till the mid-1980s, discourse might have been dominated by a labour-friendly, Keynesian position. Since then it probably has been shifting to a employer-friendly, neo-liberal position. In Project B, we wanted to grasp these shifts. Thus we decided to generate a sample that covered the period between 1964 and 2000. However, limited project time made it only possible to sample ten years from the whole time-period. Starting from 1964 we thus selected only issues from every 4th year, in order to eliminate effects of election and economic cycles (see table 2). We tried to choose a year in the middle of legislative periods in order to eliminate effects of election campaigns. This 4-year-rhythm is slightly disturbed: The 1972 and 1983 elections were predated due to political crises. We also chose these particular years to make media data more or less comparable with survey data: For the years 1984, 1991 and 2000, some ALLBUS questions measure attitudes to personal economic success, to people in need (including the unemployed) and to social security (including unemployment security). The same is true for SOEP for the years 1987, 1992, 1997 and 2002. Due to the four-year cycle, we do not have data for all, but at

least for some of these years. This rules out the possibility of identifying short-term shifts in discourse. Similar to panel data, spell effects may occur.

In contrast, Project A was interested in medium-term discourse shifts. Thus, only a six-year-period was investigated (1997 to 2002). In contrast to Project B, *all* years were chosen, thus ruling out bias concerning this dimension.

6.3 Number of Issues per Newspaper and Year

If researchers want to reconstruct historical facts (Perspective 3), they best analyse all issues of the chosen newspaper within the given time-span. This would be the case, if one used HWWA data. As stated above, this strongly limits the time-span that can be analysed.

For Perspectives 1 and 2, how many issues researchers can analyse per newspaper and year correlates negatively with the number of newspapers and years selected, too: The more issues researchers want to analyse for a given period, the fewer newspapers and shorter time-spans they can analyse. A large number of issues per year has the advantage of enabling researchers to reconstruct discursive episodes more adequately. Researchers can trace who reacted how to what. They can identify discourse communities, how arguments are twisted in discourses, how news holes are patterned and so on. If researchers want to use event history analysis or sequence analysis of short-term discourse, it is necessary to sample *all* issues from a given time-span.

Both our example research projects limit possibilities to trace short-term discourse changes: As all six years were analysed in Project A, the research team had to restrict the number of issues to three papers a week (Monday, Wednesday and Friday of each week). This made it impossible to reconstruct the thread of events and claims, and thus the interactivity of public discourses in some detail. That is, it is impossible to analyse how individual actors interact on a daily basis. Still, the sample allows to trace changes of discourse topics, opinions on a weekly basis, prominence of certain actors and so on.

In contrast, the Project B sample only allows to trace long-term discourse changes and broad tendencies because we could only sample per year. Reasons were first that data collection was more time-consuming because we actually had to collect the interviews from different archives while Project A data could be collected from the CD. Second, time and resources were a lot more limited for Project B compared to Project A.

The decision to select only a very reduced number of issues generated the problem of deciding which days to choose. A preliminary result of Project A was that coverage of unemployment issues was more or less the same all over the year (at least since 1997), with some exceptions: First, the summer time is a weak period for news coverage in general as most Germans in general and German politicians in particular are on vacation. Second, the government is usually evaluated in public after its first 100 days in power. This provides an opportunity for debates on unemployment. Finally, German discourse on unemployment is highly ritualised in the sense that unemployment figures are publicized by the central unemployment agency (BA) in regular press conference. Regularly, this gives discourse on unemployment a new impetus for a couple of days, raising the number of reports during this time. Thus we selected the day after the press conference as reference day. Starting from the reference day, we read each issue of the respective newspapers until we found at least one article on unemployment. As unemployment is highest in winter and as German business analysists usually evaluate the first quarter of the year and forecast business development for the rest of the year in March, we chose the press conference at the end of March or beginning of April.

The next step was to find our the exact dates. We know that the press conferences have been held since the 1960s. With the help of Ms. Heidelies Künzel, an BA employee, we were able to trace back the dates to 1974 (see table 3). From this information, we tried to reconstruct the ear-

lier dates: The BA disposes of the march data on the first Monday or Tuesday in April. The press conferences usually are held the following day.

First analysis of these single issues showed that they contained some but not enough information about public discourse, particularly for the earlier years. The data did not reveal whether this was due to the fact that unemployment did not constitute a contentious issue at this time, that media coverage style was more concise and officious, or whether these press conferences did not provide a strong stimulus for public debates. For these reasons, we tested different adaptive sampling strategies in order to extent the range of sampled articles.

On the one hand, we used the "Deutscher Zeitungsindex", a printed data base of news coverage by national prestige papers. This index includes bibliographic references to prominent articles about the most various issues (amongst them unemployment and labour market) for a number of years. This provided us with an easy access to substantial reports. However, the "Zeitungsindex" is available only from 1974 until 1990, an important, yet incomplete time span. Moreover, no regional newspapers and tabloids are included. Finally, selection is highly reduced and selective.

On the other hand, we had to take up the more tedious work of consulting CD-Rom data bases from 1995 onwards and going through the paper versions for the remaining years. The goal was to assemble all articles on unemployment from two latter issues. Due to this time consuming work, we have not yet finished the data gathering process for all newspapers.

The above discussion shows, that deciding on the number of newspapers, years and issues to be analysed, affect both sampling Stages 1 (selection newspapers to be analysed) and 2 (selecting and finding relevant issues). Although these three dimensions and two stages can be separated theoretically, they are inseperately intertwined in actual research process. Together, they form the first sampling phase. The second sampling phase consists of decisions on the number of articles to choose from a single issue (Stage 3) and (for Perspectives 2 and 3) the number of claims or events to choose from a single article (Stage 4). Stages 3 and 4 are intertwined as well.

6.4 Number of Articles per Issue

When deciding how many and which articles to choose from selected newspaper issues, newspaper can either draw random samples of all articles or they can read all or selected newspaper sections and sample all articles relevant to the research question. For both Projects A and B, we opted for the latter procedure. All decisions were documented in a codebook (Giugni / Statham 2002). We based these decisions on what we know about the German newspaper system and the context of German discourse on unemployment (see chapter 5).

We first defined which sections were to be analysed. Drawing on our knowledge on German newspaper formats, we consulted only the news and business sections and excluded regional and local pages for Project A. Hence, our sample reflects primarily the political news coverage of the SZ and thus the political debates within the public arena. For papers in Project B, we excluded the sports sections, leisure sections, letters to the editor, commercials and (for national papers) regional sections.

Second, a researcher read all sections and selected *all* articles, as soon as a reference was made to unemployment, irrespective of whether unemployment was the main story, a secondary topic or even a incidental reference. As we have discussed in chapter 5, the German discourse unemployment is intertwined with other discourses, especially those on social security, collective bargaining and companies' competitive behaviour. In addition, some parts of the topic "unemployment" are discussed under different labels, e. g. occupational training (youth unemployment) and immigration (ethnic differentiation of unemployment). Articles on all these topics were included into the sample. The number of articles sampled per issues was therefore rather

high. Interrelations between various policy or issue fields and discourse arenas can thus be captured within the data base.

6.5 Number of Events or Claim per Article

For Perspective 1, sampling is now finished. For Perspectives 2 and 3, researchers additionally have to decide which events and claims to sample from selected articles. The main work in sampling is collecting newspapers and choosing relevant articles. In order to choose articles, researchers have to read them completely. Thus, it is sensible to sample *all* events or claims stated in selected articles. For example, we selected all claims on unemployment (for a definition see Giugni / Statham 2002) in both research projects.

6.6 Triangulating Different Sample Types

In summary, created two different samples. For Project B, we analyzed a wide array of newspapers longitudinally. This implied an extensive sampling strategy, forcing us to reduce the amount of articles per newspaper and the number of issues per year. This strategy is the best option to study the argumentative, ideational or ideological structure of public discourses. The topic of interest are long-term changes, which can be grasped only when analysing a long period of time and a bigger number of newspapers representing various positions and arenas. In analysing the Project B sample, we assume that those arguments, ideas or ideologies dominating a public debate at a certain point of time are traceable even in single articles. This is the reason for reducing the number of texts, trusting to find at least traces of these ideas there as well.

In Project B we also analysed a number of newspapers for each point in time. Thus, we are able to catch a broader scope of the discourse by respecting different newspapers with their specific foci, topics and political orientations. While specific issue debates cannot be constructed adequately, by operating with a bigger number of newspapers, the sample is very inclusive in terms of actors, arguments and ideas. This sampling procedure is most apt for an investigation of discursive coalitions, policy communities or networks because we are able to reconstruct a wider field of claims-making actors with their particular concerns, agendas and opinions. It is possible to reconstruct these coalitions by identifying shared problem definitions, world-views and problem-solving strategies.

In contrast, we focused on only one newspaper within a shorter period of time in Project A. This allows analysing the debate very intensively by sampling a big number of articles per issue and issues per year. The decision to use one newspaper reduces the spectrum of news coverage and increases thus the data production bias. However, this might not be the primary problem from the particular research question's point of view. In fact, this kind of sample allows for an in depth analysis of issue specific debates by providing information on the semantic, syntactic and pragmatic structure of public reasoning. Moreover, we attained a more comprehensive picture of these debates by assembling articles on neighbouring issue fields (e.g., unemployment and labour markets, fiscal policies, social welfare, technological advances, international competition) and by tracing back the interrelations between various discourses (e.g., scientific, political, economic or administrative debates). Finally, we are able to reconstruct the thread of events and claims more closely. The interactivity of public discourses can thus be captured, because we may reconstruct who speaks with whom and about whom. The biased picture of public discourses, which emanates from the choice of one newspaper, is a problem, but not necessarily a decisive one, as long as we opt consciously to analyse a particular spectrum of the public arena,

e. g. the core policy domain of the most dominant political actors, onto which all prestige newspapers tend to centre anyway.

The decision to blend two sampling strategies was motivated, first, by the attempt to generate an enlarged data set that fuses the strengths of each of these two procedures and thus allows studying public discourses on unemployment more comprehensively, i.e., by amalgamating a longitudinal study of discourse arenas with a cross-sectional analysis of issue debates and fields. At the same time, we attempted to minimize the drawbacks and biases of each procedure by triangulating the samples. If many articles from one newspaper are analysed, the selected newspaper might mirror the overall public discourse well. If few articles from many newspapers are analysed, chosen articles might not be typical for the respective newspapers.

Project A provides useful information for determining how much data is necessary for a longitudinal analysis of discourse developments in Project B. Sample A suggests that broad political strands of argumentations can be uncovered within a very reduced number of articles. Yet, this is only possible if researchers apply interpretive methods of analysis. These results can be fed then into a more quantitative content analysis later on, but the first step is the more important one.

At the same time, Project B can help to determine whether the SZ, which provide the empirical basis of Project A, is typical for German discourse on unemployment. Our findings indicate that this is indeed the case. Most newspapers privilege the core policy actors to an exceptionally high extent. Newspaper differ primarily in grades of selectivity and commentaries. Unemployment initiatives are excluded from *all* selected print media. The SZ is centred on the two leading political parties and social partners, experts and think tanks. In comparison, unions are over- and liberal democrats underrepresented slightly. Yet, as Project A's data base is very large (N=2800), we have enough information about all pertinent policy actors, even some welfare organizations.

7. CONCLUSION

Process-generated data can be a valuable alternative to survey data and interviews, if the latter are not available. For some research questions, process-generated data actually are btter suited. In order to sample process-generated data, researchers first have to define the target population and cases. This definition depends on the research question. In contrast to surveys, several contrasting definitions of the target population are possible for the same data type. The reason is the multiperspectivity of process-generated data. They can be read in different ways, drawing different kinds of information from them. For example, newspapers can be used as data source for at least three different types of information: journalists' discourse behaviour, public discourse in general and historical facts. Researchers can draw most information on the first and least on the last question from newspapers. Because of this multiperspectivity and ambivalence, it is recommendable to analyse process-generated data first in detail, using interpretative methods (Hanawalt 1991). Only afterwards, researchers can decide if quantification makes sense for the particular data type.

Process-generated data are also usually biased. If and how data are biased during data production, depends first on the definition of the target population: If the newspaper or journalists' discourse behaviour are of interest, data production bias is usually no problem. For all other perspectives on the data, the way data production bias works depends on the particular combination of journalists' purposes, newspaper formats and social context. Bias varies over time, locally and depending on the particular content. Process-generated data additionally may suffer from a data selection bias: Data may be destroyed, decayed or stored inaccessibly. For newspapers, data selection is usually no problem, although researchers need to know the social context

well in order to find respective data sources. For other data types, data selection might produce additional problems. Due to biases, researchers usually cannot apply inferential statistics when using process-generated data. Instead, it is sensible to use alternative sampling strategies. In addition, data should be triangulated with other data sources in order to assess how the bias works.

Only after defining the target population and assessing bias can researchers sample process-generated data. In contrast to survey data, it is almost impossible to give general rules for sampling process-generated data. For newspapers, researchers have to decide which media to choose, which issues to choose and which articles to choose from a single issue. Thus, they can use multi-staged sampling methods. For a given sample size, researchers face a trade-off along five dimensions: (a) the number of newspapers; (b) the number of years; (c) the number of issues per year and newspaper; and (d) the number of articles per issue. Which strategy is most suitable, depends on the particular research question.

8. REFERENCES

Abbott, Andrew (1995): Things of Boundaries. *Social Research* 62. 857-882. Reprinted in: Andrew Abbott (ed.) (2001): Time Matters. On Theory and Method. Chicago: University of Chicago Press. 261-279

Altmeppen, Klaus-Dieter (2000): Funktionale Autonomie und organisationale Abhängigkeit. Inter-Relationen von Journalismus und Ökonomie. In: Löffelholz, Martin (ed.) (2000): *Theorien des Journalismus*. Wiesbaden: Westdeutscher Verlag, S. 225-239

Baglioni, Simone / della Porta, Donatella / Graziano, Paolo (2004): *Italy. National Report on WP1 (first draft) on the project: unempol. The Contentious Politics of Unemployment in Europe. Political Claim-making, Policy Deliberation and Exclusion from the Labour Market*. Report of the 'Unempol'-Project: http://ics.leeds.ac.uk/eurpolcom/unempol/papers.cfm.

Baum, Annerose / Lahusen, Christian (2004): *Germany. National Report on the project: The Contentious Politics of Unemployment in Europe. Political Claim-making, Policy Deliberation and Exclusion from the Labour Market*. Report of the 'Unempol'-Project: http://ics.leeds.ac.uk/eurpolcom/unempol/papers.cfm.

Baur, Nina (2001): *Soziologische und ökonomische Theorien der Erwerbsarbeit. Eine Einführung*. Frankfurt a. M. / New York: Campus

Baur, Nina (2004): Wo liegen die Grenzen quantitativer Längsschnittsanalysen? *Bamberger Beiträge zur empirischen Sozialforschung*. Forthcoming

Behnke, Joachim / Behnke, Nathalie / Baur, Nina (2004): *Empirische Methoden der Politikwissenschaft*. Paderborn: Ferdinand Schöningh. Forthcoming

Blossfeld, Hans-Peter / Hamerle, Alfred / Mayer, Karl Ulrich (1986): *Ereignisanalyse*. Frankfurt a. M.

Bohrmann, Hans (1999): Entwicklung der Zeitschriftenpresse. In: Wilke, Jürgen (ed.) (1999): *Mediengeschichte der Bundesrepublik Deutschland*. Bonn: Bundeszentrale für politische Bildung. S. 135-145

Bow, James (1980): The Time's Financial Markets Column in the Period Around the 1929 Crash. *Journalism Quarterly*. Volume 57, No. 2, S. 447-450

Bright, Robert / Coburn, Elaine / Faye, Julie / Gafijczuk, Derek / Hollander, Karen / Jung, Janny / Syrmbos, Helen (1999): Mainstream and Marginal Newspaper Coverage of the 1995 Quebec Referendum. An Inquiry into the Functioning of the Canadian Public Sphere. *CRSA / RCSA*. 36.3. 313-330

Brookes, R. (2004). The media representation of public opinion: British television news coverage of the 2001 general election. *Media, Culture & Society*, Vol. 26, N. 1, 63-81.

Castles, Stephen/ Kosack, Godula (1985): *Immigrant Workers and Class Structure in Western Europe*. 2nd Edition. London et al.

Chabanet, Didier / Fay, Christophe (2004): *France. National Report on the project: The Contentious Politics of Unemployment in Europe. Political Claim-making, Policy Deliberation and Exclusion from the Labour Market*. Report of the 'Unempol'-Project: http://ics.leeds.ac.uk/eurpolcom/unempol/papers.cfm.

Chiricos, Ted / Padgett, Kathy / Gertz, Marc (2000). Fear, TV news, and the reality of crime. *Criminology*, Vol. 38, N. 3, 755-787.Preliminary Analysis – France (1995 – 2002). National Report (WP1) on the project: unempol. The

Contentious Politics of Unemployment in Europe. Political Claim-making, Policy Deliberation and Exclusion from the Labour Market

Cochran, William G. (1972): *Stichprobenverfahren*. Berlin / New York: Walter de Gruyter

Creswell, John W. (1998): *Qualitative Inquiry and Research Design. Choosing Among Five Traditions*. Thousand Oaks / London / New Delhi: Sage.

Esping-Andersen, Gøsta (1990): *The Three Worlds of Welfare Capitalism*. Cambridge (UK) / Oxford (UK): Polity / Blackwell

Flick, Uwe (2000): Triangulation in der qualitativen Forschung. In: Flick, Uwe / Kardoff, Ernst von / Steinke, Ines (ed.) (2000): *Qualitative Sozialforschung. Ein Handbuch*. Reinbek: Rowohlt Taschenbuch Verlag. 309-318

Franzosi, Roberto (1987): The Press as a Source of Socio-Historical Data: Issues in the Methodology of Data Collection from Newspapers. *Historical Methods*, *20*, 5-16.

Galtung, J. / Ruge, M. (1973): Structuring and selecting news. In S. Cohen & J. Young (Eds.), *The Manufacture of News: Social Problems, Deviance and the Mass Media* (pp. 62-72). London: Constable.

German Information Center (1995): *Citizenship and Naturalisation*. http://langlab.uta.edu/GERM/GIC/focus/FOCUS_95_2_January.txt (February 1995; 950 Third Avenue. New York)

Gesterkamp, Thomas (1993): So selbstverständlich und doch so unbekannt. Die Arbeitswelt in den Medien. *Medium*. Volume 23, No.4, 4-6

Gigerenzer, Gerd (1999): Über den mechanischen Umgang mit statistischen Methoden. In: Roth, Erwin / Holling, Heinz (ed.) (1999): *Sozialwissenschaftliche Methoden. Lehr- und Handbuch für Forschung und Praxis*. 5.Auflage. München / Wien: R. Oldenbourg. S. 607-618

Giugni, Marco / Statham, Paul (2002): *The Contentious Politics of Unemployment in Europe. Political Claim-Making, Policy Deliberation and Exclusion from the Labor Market. Codebook Workpackage 1: Political Claim-Making in the Public Domain*. Reihe: European Political Communication Working Paper Series. Volume 2/02. LeedsGenua: http://ics.leeds.ac.uk/eurpolcom/unempol/papers.cfm.University of Geneva. http://www.leeds.ac.uk/ics/euro/unempol

Gleich, Uli (1998): Politikvermittlung und politische Partizipation durch Medien. In: Dichanz, Horst (ed.) (1998): *Handbuch Medien: Medienforschung. Konzepte, Themen, Ergebnisse*. Bonn: Bundeszentrale für politische Bildung. S. 54-59.

Groß, Rolf (1996): Zur Konzentration auf dem Gebiet der Printmedien. *Zeitschrift für Urheber- und Medienrecht. ZUM.* Volume 40, No. 5, 365-378

Hagen, Lutz M. (1996): Die großen internationalen Medienkonzerne. In: Mast, Claudia (ed.) (1996): *Markt – Macht – Medien*. Konstanz: UVK Medien, 119-130

Hanawalt, Barbara A. (1991): The Voices and Audiences of Social History Records. *Social Science History*. Volume 15. No. 2. 159-175

Hell-Berlin, Katja Kristina (1997): Das Land der unbegrenzten kommerziellen Möglichkeiten. Zur Entwicklung der Medien in den USA. In: Rupp, Hans Karl / Hecker, Wolfgang (ed.) (1997): *Auf dem Weg zur Telekratie? Perspektiven der Mediengesellschaft*. Konstanz: UVK Medien Verl.-Ges., 70-94

Hocke, Peter (1998): Determining the Selection Bias in Local and National Newspaper Reports on Protest Events. In: Rucht, Dieter / Koopmans, Ruud / Neidhardt, Friedhelm (ed.) (1998): *Acts of Dissent*. Berlin: Edition Sigma

Hoffmann, Jochen / Sarcinelli, Ulrich (1999a): Politische Wirkungen der Medien. In: Wilke, Jürgen (ed.) (1999): *Mediengeschichte der Bundesrepublik Deutschland. Bonn: Bundeszentrale für politische Bildung*. S. 720-748

Jansen, Andrea / Ruberto, Rosaia / Münch, Richard (1997): *Mediale Konstruktion politischer Realität. Politikvermittlung im Zeitalter der Fernsehdemokratie*. Wiesbaden: Dt. Univ.-Verl.

Jarren, Otfried (1996): Publizistische Märkte und Kommunikationspolitik. Öffentliche Regulierung statt politisch-administrativer Steuerung. In: Altmeppen, Klaus-Dieter (ed.) (1996): *Ökonomie der Medien und des Mediensystems. Grundlagen, Ergebnisse und Perspektiven medienökonomischer Forschung*. Opladen: Westdeutscher Verlag, 203-219

Jarren, Otfried / Meier, Werner A. (2002): Mediensysteme und Medienorganisation als Rahmenbedingungen für den Journalismus. In: Jarren, Otfried / Weßler, Hartmut (ed.) (2002): *Jouralismus – Medien – Öffentlichkeit*. Wiesbaden: Westdeutscher Verlag, 99-163

Keppler, Angela (2000): Medien und Kommunikationssoziologie. Verschränkte Gegenwarten. Die Untersuchung kultureller Transformationen. *Soziologische Revue*. Volume 23, Sonderheft 5, 140-153

Kepplinger, Hans Matthias (1994): Publizistische Konflikte. Begriffe, Ansätze, Ergebnisse. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*. Sonderheft 24. Vol. 46. 214-233

Kepplinger, Hans Mathias (1999a): Zeitungsberichterstattung im Wandel. In: Wilke, Jürgen (ed.) (1999*): Mediengeschichte der Bundesrepublik Deutschland*. Bonn: Bundeszentrale für politische Bildung. S. 195-210

Kepplinger, Hans Mathias (1999b): Publizistische Konflikte. In: Wilke, Jürgen (ed.) (1999): *Mediengeschichte der Bundesrepublik Deutschland*. Bonn: Bundeszentrale für politische Bildung. S. 698-719

Kiefer, Marie-Luise (1998): Massenkommunikation 1995. Ergebnisse der siebten Welle der Langzeitstudie zur Mediennutzung und Medienbewertung. In: Dichanz, Horst (ed.) (1998): *Handbuch Medien: Medienforschung. Konzepte, Themen, Ergebnisse*. Bonn: Bundeszentrale für politische Bildung. S. 20-37

Kiefer, Markus (2000): Objektive, wahrhaftige und vollständige Berichterstattung. Punktuelle Beobachtungen zu einer zunehmenden Außensteuerung des Journalismus. In: Mahle, Walter A. (ed.) (2000): *Orientierung in der Informationsgesellschaft*. Konstanz: UVK Medien Verl.-Ges., S. 82-83

Kleinsteuber, Hans J. / Thomaß, Barbara (1996): Konkurrenz versus Konzentration. Unternehmensverflechtungen im globalen Medienmarkt. In: Meckel, Miriam / Kriener, Markus (ed.) (1996): *Internationale Kommunikation*. Opladen: Westdeutscher Verlag, -144

Kuenheim, Eberhard von (1996): Verantwortung der Industrie – Verantwortung der Medien. Analogie oder Widerspruch. In: Hamm, Ingrid (ed.) (1996): *Verantwortung im freien Medienmarkt. Internationale Perspektiven zur Wahrung professioneller Standards*. Gütersloh: Verl. Bertelsmann Stiftung, 61-70

Küng, Lucy (2001): The Internet's impact on incumbent media firms. A management perspective. *Medien & Kommunikationswissenschaft*. Volume 49, No. 2, 218-226

Lahusen, Christian / Baur, Nina (2004): *Do Media Discourses Influence Opinions? Connecting Newspaper Discourse on Unemployment with Attitude Change on Social Security*. Paper Presented on the RC 33 Sixth International Conference on Social Science Methodology: Recent Developments and Applications in Social Research Methodology. 16 – 20. August 2004, Amsterdam

Lerg, Winfried B. / Schmolke, Michael (1995): Die Zeitung als Quelle. Das Auswahlproblem. *Relation. Medien – Gesellschaft – Geschichte*. Volume 2. No. 1. 11-18

Linders, Annulla (2004): *Sweden. National Report on the project: The Contentious Politics of Unemployment in Europe. Political Claim-making, Policy Deliberation and Exclusion from the Labour Market*. Report of the 'Unempol'-Project: http://ics.leeds.ac.uk/eurpolcom/unempol/papers.cfm.

MacDevitt, Michael (2003): In defense of autonomy. A critique of the public journalism critique. *Journal of Communication*, Vol. 53, N. 1, 155-165.Political Claim-making in the Public Domain (WP1). National Report Sweden. Draft. Project: unempol. The Contentious Politics of Unemployment in Europe. Political Claim-making, Policy Deliberation and Exclusion from the Labour Market

MacQuail, Denis (2003). *Mass Communication Theory*. London: Sage, 4th ed.

Mayer, Karl Ulrich (1998): Causality, Comparisons and Bad Practices in Empirical Social Research. A Comment on Stanley Lieberson's Chapter. In: Blossfeld, Hans-Peter / Prein, Gerald (ed.) (1998): *Rational Choice Theory and Large-Scale Data Analysis*. Boulder: Westview. S. 146-157

McCarthy, John D. / McPhail, Clark / Smith, Jackie (1996): Images of Protest. Dimensions of Selection Bias in Media Coverage of Washington Demonstrations, 1982 and 1991. *American Sociological Review*. Volume 61. 478-499

Müller, Albert (1996): Die Zeitung. Eine Quelle der Historischen Sozialwissenschaften. *Relation*. No. 1 3.1996. Vol. 3. 45-48

Münch, Richard (1993): Journalismus in der Kommunikationsgesellschaft. *Vierteljahreshefte für Kommunikationsforschung. Zeitschrift für die Wissenschaft von Presse, Rundfunk, Film, Rhetorik, Öffentlichkeitsarbeit*. Volume 38, No. .3. 261-279

Münch, Richard (1997a): Mediale Ereignisproduktion. Strukturwandel der politischen Macht. In: Hradil, Stefan (ed.) (1997): *Differenz und Integration. Die Zukunft moderner Gesellschaften. Verhandlungen der der Deutschen Gesellschaft für Soziologie in Dresden 1996*. Frankfurt am Main: Campus

Münch, Richard (1997b): Mediale Kommunikationsdynamik. In: Schanze, Helmut / Ludes, Peter (ed.) (1997): *Qualitative Perspektiven des Medienwandels. Positionen der Medienwissenschaft im Kontext „Neuer Medien"*. Opladen: Westdeutscher Verlag, S.77-79

Oberst-Hundt, Christina / Oberst, Walter (2001): Politik – Wirtschaft – Medien. Rückblick auf die Entwicklung nach 1945. *Medien und Erziehung*. Volume 45, No. 4, 223-228

Oliver, Pamela E. / Maney, Gregory M. (2000): Political Processes and Local Newspaper Coverage of Protest Events: From Selection Bias to Triadic Interactions. *American Journal of Sociology*. Volume 106. No. 2.

Oliver, Pamela E. / Myers, Daniel J. (1998): How Events Enter the Public Sphere. Conflict, Location and Sponsorship in Local Newspaper Coverage of Public Events. *American Journal of Sociology*. Volume 105. 38-87

Ostner, Ilona (1995): Arm ohne Ehemann? Sozialpolitische Regulierung von Lebenschancen für Frauen im internationalen Vergleich. *Aus Politik und Zeitgeschichte.* B 36-37. 3-12

Pietilä, Veikko (1992): Beyond the News Story: News as Discursive Composition. *European Journal of Communication*. Volume 7. No. 1. 37.68

Pöttker, Horst (2002): Wann dürfen Journalisten Türken Türken nennen? Zu Aufgaben und Systematik der Berufsethik am Beispiel des Diskriminierungsverbots. *Publizistik. Vierteljahreshefte für Kommunikationsforschung*, Vol. 47, N. 3, 265-280.

Röper, Horst / Pätzold, Ulrich (1993): *Medienkonzentration in Deutschland. Medienverflechtungen und Branchenvernetzungen*. Berlin: Freie Universität Berlin, Fachinformationsstelle Publizistik

Rupp, Hans Karl (1997): Die politische Macht des Fernsehens und Vorschläge, sie zu begrenzen. In: Rupp, Hans Karl / Hecker, Wolfgang (ed.) (1997): *Auf dem Weg zur Telekratie? Perspektiven der Mediengesellschaft*. Konstanz: UVK Medien Verl.-Ges., 223-245

Schmitt-Beck, Rüdiger / Pfetsch, Barbara (1994): Politische Akteure und die Medien der Massenkommunikation. Zur Generierung der Öffentlichkeit in Wahlkämpfen. *Kölner Zeitschrift für Soziologie und Sozialpsychologie*. Sonderheft 34. Volume 46. 106-138

Schnell, Rainer (1986): *Missing-Data-Probleme in der empirischen Sozialforschung*. Inaugural-Dissertation zur Erlangung des akademischen Grades eines Doktors der Sozialwissenschaft an der Ruhr-Universität Bochum – Abteilung Sozialwissenschaft.

Schnell, Rainer (1997): *Nonresponse in Bevölkerungsumfragen. Ausmaß, Entwicklung und Ursachen*. Opladen: Leske + Budrich

Schulz, Rüdiger (1999): Nutzung von Zeitungen und Zeitschriften. In: Wilke, Jürgen (ed.) (1999): *Mediengeschichte der Bundesrepublik Deutschland*. Bonn: Bundeszentrale für politische Bildung. S. 401-425

Schütz, Walter J. (1999): Entwicklung der Tagespresse. In: Wilke, Jürgen (ed.) (1999): *Mediengeschichte der Bundesrepublik Deutschland*. Bonn: Bundeszentrale für politische Bildung. S. 109-134

Seale, Clive (1999): *The Quality of Qualitative Research*. London / Thousand Oaks / New Delhi: Sage

Statham, Paul (2004): Great Britain. National Report on the project: The Contentious Politics of Unemployment in Europe. Political Claim-making, Policy Deliberation and Exclusion from the Labour Market. Report of the 'Unmepol'-Project: . http://ics.leeds.ac.uk/eurpolcom/unempol/papers.cfm

Steinhage, Nikolei (2000): Zeitaggregation und Ereignisdaten. Eine Simulation zu den Auswirkungen der Prozeßzeitskalierung. *Globalife Working Paper* No. 3 / 2000. Fakultät für Soziologie an der Universität Bielefeld. Bielefeld

Trappel, Josef / Meier, Werner A. / Schrapa, Klaus / Wölk, Michaela (2002): *Die gesellschaftlichen Folgen der Medienkonzentration. Veränderungen in den demokratischen und kulturellen Grundlagen der Gesellschaft*. Opladen: Leske und Budrich

Tudor, Andrew (1992): Them and us. Story and stereotype in TV world cup coverage. *European journal of communication*, No. 3, Vol. 7, S. 391-413.

van den Berg, Harry / Smit, Johannes H. / van der Veer, Kees (1992): Contextualization in Newspaper Articles: A Sequential Analysis of Actors' Quotes on the PACTO Affair. *European Journal of Communication*. Vol. 7. 359-389

van Dijk, Teun A. (1988): *News Analysis. Case Studies of International and National News in the Press.* Hillsdale, N.J.: Lawrence Erlbaum Ass.

van Dijk, Teun A. (ed.). (1997): *Discourse as Social Interaction. Discourse Studies: A Multidisciplinary Introduction.* London: Sage.

Wilke, Jürgen (1999a): Überblick und Phasengliederung. In: Wilke, Jürgen (ed.) (1999): *Mediengeschichte der Bundesrepublik Deutschland*. Bonn: Bundeszentrale für politische Bildung. 15-27, 777-815

Wilke, Jürgen (1999b): Leitmedien und Zielgruppenorgane. In: Wilke, Jürgen (ed.) (1999): *Mediengeschichte der Bundesrepublik Deutschland*. Bonn: Bundeszentrale für politische Bildung. 302-329

Wittenzellner, Helmut (ed.) (2000): *Internationalisierung der Medienindustrie. Entwicklung, Erfolgsfaktoren und Handlungsempfehlungen*. Stuttgart: LOG_X Verlag

Wolf, Michael J. (1999): *The Entertainment economy. How mega-media forces are transforming our lives*. New York: Times Books

Wolff, Volker (2002): Wer ist das Futtertier? Nachrichtenproduktion. *Message*. No. 1, S. 78-81

Zakrzewski, Raimund H. (1995): Marketingforschung für eine Tageszeitung. Primär- und Sekundärerhebungen der Süddeutschen Zeitung. In: Böhme-Dürr, Karin / Graf, Gerhard (ed.) (1995): *Auf der Suche nach dem Publikum. Medienforschung für die Praxis*. Konstanz: UVK. S. 45-67.

Zerdick, Axel (1994): Ist die Unabhängigkeit der Medien bedroht? Neue Entwicklung der Medienkonzentration in Europa. *Medium*. Volume 24, No. 1, 11-14.